PCT WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6 : G08B 13/196

A1

(11) Internationale Veröffentlichungsnummer: **WO 97/04428**

(43) Internationales Veröffentlichungsdatum: 6. Februar 1997 (06.02.97)

(21) Internationales Aktenzeichen: PCT/DE96/01344

(22) Internationales Anmeldedatum: 22. Juli 1996 (22.07.96)

(30) Prioritätsdaten:

195 26 508.4	20. Juli 1995 (20.07.95)	DE
196 00 958.8	12. Januar 1996 (12.01.96)	DE

(71) Anmelder *(für alle Bestimmungsstaaten ausser US)*: FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. [DE/DE]; Leonrodstrasse 54, D-80636 München (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder *(nur für US)*: WAGNER, Thomas [DE/DE]; Spinnereistrasse 20, D-91052 Erlangen (DE). FRISCHHOLZ, Robert [DE/DE]; Noetherstrasse 12, D-91058 Erlangen (DE). BAUER, Norbert [DE/DE]; Volckamerstrasse 14, D-91058 Erlangen (DE).

(74) Anwalt: OLGEMÖLLER, Luitgard; Leonhard, Olgemöller, Fricke, Josephspitalstrasse 7, D-80331 München (DE).

(81) Bestimmungsstaaten: US, europäisches Patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Veröffentlicht

Mit internationalem Recherchenbericht.

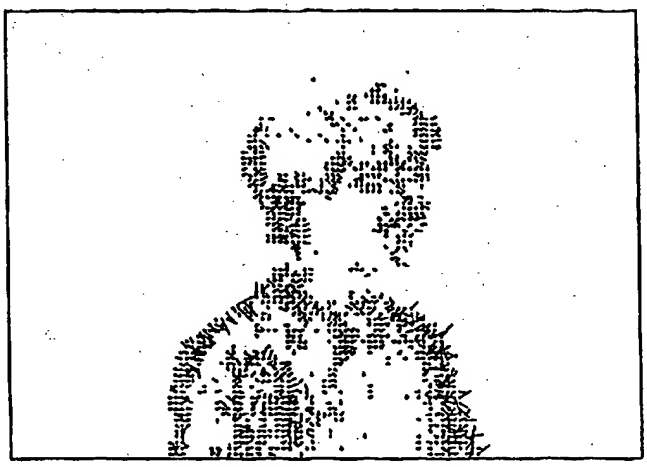
Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist. Veröffentlichung wird wiederholt falls Änderungen eintreffen.

(54) Title: INTERACTIVE SURVEILLANCE SYSTEM

(54) Bezeichnung: INTERAKTIVES ÜBERWACHUNGSSYSTEM

(57) Abstract

A process is disclosed for observing or monitoring an area in space. Images of the area (1) to be observed/monitored are taken by means of at least one camera (2), the converted image information is sent to a control computer (3), the control computer (3) checks the images for changes in the observed area (1), the main direction and/or speed of optically perceptible changes (movements) in the observed area are analysed by analysis of the optical flow, the analysis result is converted into control instructions sent to a controller (4) that mechanically moves the camera, and if required the position and/or size of the image section that may be taken by the camera is modified. Also disclosed is an interactive surveillance system with a control computer (3) connected to or integrated into a camera (2) or into one among several cameras. The control computer (3) receives and processes the data recorded by the camera(s) and sends control signals to a zoom lens (5). A mechanical controller (4) reacts to impulses from the control computer (3) and modifies the recording section of the area (1) to be observed or the size of the image section (5). A device reacts to signals from the control computer (3) and opens a normally closed passage.



(57) Zusammenfassung

Die Erfindung betrifft ein Verfahren zum Beobachten oder Überwachen eines räumlichen Bereiches, umfassend das Aufnehmen von Bildern aus dem zu beobachtenden/überwachenden Bereich (1) mit Hilfe mindestens einer Kamera (2), das Senden der umgewandelten Bildinformation an einen Steuerrechner (3), das Auswerten der Bilder im Steuerrechner (3) in bezug auf Veränderungen im zu überwachenden Bereich (1), das Analysieren der Hauptrichtung und/oder der Geschwindigkeit von optisch wahrnehmbaren Veränderungen (Bewegungen) im zu überwachenden Bereich mit Hilfe der Analyse des optischen Flusses, das Umsetzen des Analyseergebnisses in Steuerbefehle zum mechanischen Bewegen der Kamera an eine Steuereinrichtung (4), und ggf. das Verändern des von der Kamera aufnehmbaren Bildausschnittes in bezug auf Lage und/oder Größe; des weiteren betrifft die Erfindung ein interaktives Überwachungssystem, welches einen mit einer Kamera (2) oder einer von mehreren Kameras (2) verbundenen oder darin integrierten Steuerrechner (3), der die von der/den Kamera(s) aufgenommenen Daten aufnimmt und verarbeitet und Steuersignale an ein Zoom-Objektiv (5) senden kann, ferner eine mechanische Steuereinrichtung (4), die auf Impulse des Steuerrechners (3) hin den Aufnahmeausschnitt aus dem zu überwachenden Bereich (1) oder die Größe des Bildausschnittes (5) verändern kann, sowie eine Einrichtung zum Öffnen einer üblicherweise verschlossenen Passage aufweist, die auf Signale des Steuerrechners (3) reagiert.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AM	Armenien	GB	Vereinigtes Königreich	MX	Mexiko
AT	Österreich	GE	Georgien	NE	Niger
AU	Australien	GN	Guinea	NL	Niederlande
BB	Barbados	GR	Griechenland	NO	Norwegen
BE	Belgien	HU	Ungarn	NZ	Neuseeland
BF	Burkina Faso	IE	Irland	PL	Polen
BG	Bulgarien	IT	Italien	PT	Portugal
BJ	Benin	JP	Japan	RO	Rumänien
BR	Brasilien	KE	Kenya	RU	Russische Föderation
BY	Belarus	KG	Kirgisistan	SD	Sudan
CA	Kanada	KP	Demokratische Volksrepublik Korea	SE	Schweden
CF	Zentrale Afrikanische Republik	KR	Republik Korea	SG	Singapur
CG	Kongo	KZ	Kasachstan	SI	Slowenien
CH	Schweiz	LI	Liechtenstein	SK	Slowakei
CI	Côte d'Ivoire	LK	Sri Lanka	SN	Senegal
CM	Kamerun	LR	Liberia	SZ	Swasiland
CN	China	LK	Litauen	TD	Tschad
CS	Tschechoslowakei	LU	Luxemburg	TG	Togo
CZ	Tschechische Republik	LV	Lettland	TJ	Tadschikistan
DE	Deutschland	MC	Monaco	TT	Trinidad und Tobago
DK	Dänemark	MD	Republik Moldau	UA	Ukraine
EE	Estland	MG	Madagaskar	UG	Uganda
ES	Spanien	ML	Mali	US	Vereinigte Staaten von Amerika
FI	Finnland	MN	Mongolei	UZ	Usbekistan
FR	Frankreich	MR	Mauretanien	VN	Vietnam
GA	Gabon	MW	Malawi		

Interaktives Überwachungssystem

Die vorliegende Erfindung betrifft ein interaktives Beobachtungs- oder Überwachungssystem sowie ein Verfahren zum Überwachen bzw. Beobachten eines räumlichen Bereiches, insbesondere zur Personenüberwachung oder Identifizierung von Personen.

Der Stand der Technik kennt Kamerasysteme für Überwachungsaufgaben. Diese Kamerasysteme beschränken sich jedoch auf reine Visualisierungs- und Übertragungsaufgaben, d.h. die Bildinformationen aus dem zu überwachenden Bereich werden zur direkten, gleichzeitigen Beobachtung an Monitore in einer Zentrale übermittelt, oder sie werden auf Videoband abgespeichert, um später abgerufen werden zu können, bzw. sie werden nach einiger Zeit automatisch gelöscht.

Aus der EP 90 104 469.3 ist ein Verfahren zur Detektion von Intrudern bekannt, in welchem mit einer Videokamera und einer Video-Bild-Verarbeitungseinheit ein vorgegebener Überwachungsbereich überwacht wird. Dabei werden die zu detektierenden Objekte durch zwei zeitlich versetzte Videobilder mit Hilfe eines Differenzbildverfahrens erkannt.

Prinzipiell werden im Differenzbildverfahren aufeinanderfolgende Bilder voneinander abgezogen. Tritt eine hinreichend große Differenz zwischen zwei Bildern auf, so wird diese Differenz als "Bewegung" identifiziert bzw. analysiert.

Der Einsatz solcher Verfahren hat jedoch Nachteile. So wird die Gesamtbewegungsrichtung im Falle komplexer Strukturen und einer inhomogenen Bewegungen (Beispiel: ein Arm schlenkert gegen die Laufrichtung) häufig inkorrekt ermittelt, da die Bewegungsrichtung anhand des ermittelten Schwerpunktes des zu beobachtenden Objekts analysiert wird.

Aufgabe der vorliegenden Erfindung ist es, ein Überwachungssystem bereitzustellen, in welchem die Kamera in Abhängigkeit vom Bildinhalt automatisch gesteuert und nachgeführt werden kann, wobei die Gesamtbewegungsrichtung mit hoher Genauigkeit ermittelt wird.

Gelöst wird die vorgenannte Aufgabe mit Hilfe des Verfahrens gemäß Anspruch 1.

Das interaktive Überwachungssystem ist eine "intelligente Kamera", die selbsttätig und schnell auf durch eine Analyse des optischen Flusses gewonnene Bildsignale reagieren kann. Es handelt sich dabei um eine Echtzeit-Bildauswertung einer oder mehrerer Kameras, die der Bewegung in einem Überwachungsbereich nachgeführt werden kann bzw. können und/oder über die Veränderung des Bildausschnittes (Zoom) eine Vergrößerung interessierender Bereiche bewirkt und somit z.B. die Gesichter von Menschen herausvergrößern kann. Das System eignet sich für die Überwachung von Gebäuden (Eingängen, Tiefgaragen) und dgl.

Figur 1a-c zeigt die Auswertung von zu verschiedenen Zeiten (t und t+1) aufgenommenen Bildern (Fig. 1a) und zwar mit Hilfe der Analyse des optischen Flusses (Fig. 1b) bzw. mit Hilfe eines Differenzbildverfahrens (Fig. 1c). In Figur 1b erkennt man ein optisches Flußvektorfeld. Das Kreuz zeigt jeweils den Ursprung der Bewegungsrichtung, Länge und Richtung des jeweiligen Striches weisen Länge und Richtung des Bewegungsvektors.

Durch z.B. Vektorsummation aller Einzelrichtungen kann nun auf eine Gesamtbewegungsrichtung von Objekten im Bild geschlossen und somit die Kamera nachgeführt werden. Mit einem reinen Differenzbild ist dies, wie aus Figur 1c erkennbar, nicht möglich.

Figur 2 zeigt ein System mit einer Autofocus-Kamera 2 mit einem Zoom-Objektiv 5, die auf einen Überwachungsbereich 1 gerichtet ist. Die Kamera gibt die aufgenommenen Bildsignale an einen Steuerrechner 3 weiter, der wiederum mit einer mechanischen Steuereinrichtung 4 verbunden ist. Die Kamera kann in verschiedener Weise einer ggf. detektierten Bewegung nachgehen, beispielsweise indem sie die Bewegungsrichtung analysiert und dieser Richtung im zu überwachenden Bereich folgt oder indem sie eine Vergrößerung eines interessierenden Bildausschnittes durch Zoomen bewirkt.

Die Kamera kann kontinuierlich Bilder vom zu überwachenden Bereich an den Steuerrechner senden. Selbstverständlich sind auch intermittierende Aufnahmen oder ein diskontinuierliches Aufnehmen (nur zu interessierenden, vorherbestimmten oder nicht vorherbestimmten Zeitabschnitten) möglich. Im Anfangszustand ist dabei das an der Kamera angeschlossene Zoom-Objektiv auf den größtmöglichen Blickwinkel eingestellt.

Die Bilder werden vom Steuerrechner (3) kontinuierlich ausgewertet, wobei nach Bewegung im Bild gesucht wird. Tritt Bewegung im Bild auf, so wird mit Hilfe einer Analyse des optischen Flusses die Hauptrichtung der Bewegung und deren Geschwindigkeit ermittelt. Durch die Vektorsummation aller

Einzelrichtungen wird dabei auf eine Gesamtbewegungsrichtung von Objekten im Bild geschlossen, worauf die Kamera geeignet nachgeführt werden kann, so daß das bewegte Objekt immer voll im Bild ist.

Der Steuerrechner ermittelt aus dem Ort der Bewegung im Bild Steuerbefehle, um die Kamera auf eine geeignete Vergrößerung der Bewegung nachzustellen. Dabei werden die Bewegungen bezüglich des "Blickwinkels" der Kamera, also ihrer Lage bzw. des Ausschnittes aus dem zu überwachenden Bereich, durch eine mechanische Steuereinrichtung bewirkt, die ihre Impulse vom Steuerrechner erhält. Die Impulse zum Verändern des Zoom-Ausschnittes der Kamera werden vom Steuerrechner an die Kamera zurückgeleitet. Durch die Kombination dieser beiden "intelligenten" Maßnahmen wird die Ursache der Bewegung (die Person, das Tier, das sich bewegt) optimal ins Bild gebracht und vergrößert, so daß z.B. die Bewegungen einer Person vollständig von der Kamera aufgenommen werden, solange sie sich im zu überwachenden Bereich aufhält.

Parallel dazu kann mit einem Mustervergleichsverfahren, wie es z.B. aus dem Stand der Technik bekannt ist, im Bild nach gesichtsfarbigen Objekten gesucht werden und speziell ein Gesicht einer Person im Überwachungsbereich herausvergrößert werden. So kann ein derartiges Mustervergleichsverfahren z.B. nach ellipsenförmigen Objekten bestimmter Ausrichtung, Größe und Helligkeit suchen, die im Normalfall Gesichter darstellen.

Vorzugsweise schaltet die Kamera in regelmäßigen Abständen auf den günstigsten Blickwinkel und den ihr größtmöglichen Ausschnitt zurück, um evtl. weitere Bewegungen im Überwachungsbereich zu detektieren.

Die interaktive Kamera kann auch mit existierenden Gesichts- und Personenerkennungssystemen gekoppelt werden, um Aussagen über die Identität der im Überwachungsbereich befindlichen Person zu ermitteln. Hierzu wird z.B. ein Ausschnitt, der anhand der Bewegungsanalyse als zu einem bewegten Gesicht gehörig erkannt

wurde, herausvergrößert und dieses Teilbild einem Mustervergleich zugeführt, der in einer Datenbank mit abgespeicherten Gesichtsbildern nach dem Namen der entsprechenden Person sucht.

In einer besonderen Ausgestaltung des erfindungsgemäßen Verfahrens kann durch automatischen Mustervergleich des aufgenommenen Bewegungsmusters mit abgespeicherten Mustern zwischen Personen und anderen bewegten Objekten (Blätter im Wind, Tieren usw.) unterschieden werden, um Fehlalarme zu vermeiden.

Selbstverständlich kann das interaktive Überwachungssystem nicht nur mit einer Kamera, wie oben beschrieben, ausgestattet sein. Bei der Verwendung mehrerer Kameras können Stereobildauswertungsverfahren zur Detektion und Erkennung von bewegten Objekten herangezogen werden.

Das vorstehend beschriebene interaktive System zur Beobachtung bzw. Überwachung von Bereichen wie Gebäuden (Eingängen, Tiefgaragen) oder dgl. führt durch eine Echtzeit-Bildauswertung die Kamera oder die Kameras der Bewegung nach und vergrößert den bewegten Gegenstand optimal heraus. Der bewegte Gegenstand, z.B. eine menschliche Person, kann automatisch klassifiziert/identifiziert werden, wobei sich Fehlalarme durch die Bewegung von Tieren, Blättern usw. vermeiden lassen. In der Kombination mit bekannten Bild- und Signalverarbeitungsverfahren eignet sich das Verfahren und das System der vorliegenden Erfindung zur namentlichen Identifikation der im Überwachungsbereich befindlichen Personen und dadurch auch als Eingangskontrollsystem für Bereiche, die nur einer definierten Personengruppe zugänglich sein dürfen (Ansprüche 8 und 12).

Ansprüche:

1. Verfahren zum Beobachten oder Überwachen eines räumlichen Bereiches, umfassend
   - das Aufnehmen von Bildern aus dem zu beobachtenden/überwachenden Bereich (1) mit Hilfe mindestens einer Kamera (2),
   - das Senden (2a) der umgewandelten Bildinformation (A/D) an einen Steuerrechner (3),
   - das Auswerten der Bilder im Steuerrechner (3), um eine erste Veränderung im zu überwachenden Bereich (1) zu erfassen,
   - das vektorielle Ermitteln der Bewegungsrichtung der optisch wahrgenommenen ersten Veränderung (Bewegungen) im zu überwachenden Bereich mit Hilfe des optischen Flusses,
   - das Umsetzen des Ermittlungsergebnisses in Steuerbefehle an eine Steuereinrichtung (4) zum mechanischen Bewegen der Kamera, und ggf.
   - das Verändern (5) des von der Kamera erfaßten Bildausschnittes in Bezug auf Größe.

2. Verfahren nach Anspruch 1, worin beim Erkennen von Bewegung im zu überwachenden Bereich optische, akustische oder elektrische (Alarm-) Signale abgegeben werden.

3. Verfahren nach einem der obigen Ansprüche, worin der Bildausschnitt mindestens einer Kamera (2) beim Erkennen von Bewegung durch Verändern eines Zoom-Objektes (5) vergrößert wird.

4. Verfahren nach einem der voranstehenden Ansprüche, wobei zusätzlich parallel mit einem Mustervergleichsverfahren im räumlichen Bereich nach gesichtsförmigen Objekten gesucht wird und bei Erkennen des sich bewegenden Objektes als Gesicht dieses mit Hilfe eines Zooms (5) der Kamera (2) vergrößert und/oder seiner Bewegung mit der Kamera (2) nachgeführt wird.

5. Verfahren nach einem der voranstehenden Ansprüche, worin in vorbestimmten Abständen die ursprüngliche Lage und/oder der ursprüngliche Bildausschnitt der Kamera (2) wiederhergestellt werden, um eine zweite oder dritte Veränderung im räumlichen Bereich (ebenfalls) zu erfassen.

6. Verfahren nach einem der obigen Ansprüche, worin das Bilder-Aufnehmen kontinuierlich oder intermittierend ist.

7. Verfahren nach einem der voranstehenden Ansprüche, wobei dem Steuerrechner (3) zusätzlich Informationen über Gesichter oder Personen aus einem Personenidentifikationsverfahren zugänglich sind.

8. Verfahren nach Anspruch 7, worin die Bildinformationen der Kamera (2) mit gespeicherten Informationen zu definierten Personen verglichen wird und der Steuerrechner (3) mit einer Einrichtung verbunden ist, mit deren Hilfe beim Feststellen des Vorhandenseins von ausreichender Übereinstimmung dahingehend, daß Personenidentität angenommen wird, eine üblicherweise verschlossene Passage geöffnet werden kann.

9. Verfahren nach einem der voranstehenden Ansprüche, wobei analysierte Bewegungsabläufe mit den Signalen bereits gespeicherter Bewegungsabläufe verglichen werden und die Abgabe von (Alarm-) Signalen nur dann erfolgt, wenn der Vergleich deutliche Abweichungen ergibt.

10. Verfahren nach Anspruch 9, wobei der Bewegungsvergleich kontinuierlich oder intermittierend erfolgt.

11. Verfahren nach einem der voranstehenden Ansprüche, worin die auszuwertenden Bilder Stereobilder sind, die von insbesondere zwei Kameras gleichzeitig aufgenommen werden, um der Haupt-Bewegungsrichtung (v1,v2) der ersten und der zweiten "Veränderung" jeweils eine erste (v1',v2') und eine zweite (v1",v2") - senkrecht zueinander orientierte - Bewegeungskomponente zuzuweisen.

12. Verfahren nach einem der voranstehenden Ansprüche, worin nach zyklischem Wiederherstellen des ursprünglichen (größtmöglichen) Bilsausschnitts einer der zwei Kameras (2), die im Bildausschnitt zurückgestellte Kamera eine "Veränderung" als solches ermittelt und die nicht zurückgestellte Kamera der Bewegung der schon erfaßten "Veränderung" weiter nachgeführt wird.

13. Verfahren nach einem der voranstehenden Ansprüche, worin die zyklische im Bildausschnitt zurückgestellte Kamera nach vorbestimmter Zeit der Nicht-Erfassung einer weiteren "Veränderung" auf die von der anderen Kamera bereits erfaßte (erste) "Veränderung" schnell zielgeführt wird, um mit beiden Kameras mehrere Bewegungskomponenten (v1,v2; v1',v2'; v1",v2") oder ein Stereobild zu erfassen.

14. Interaktives Überwachungssystem, umfassend
- mindestens eine Kamera (2) mit einem Zoom-Objektiv (5),
- einen mit einer Kamera oder einer der Kameras verbundenen oder darin integrierten Steuerrechner (3), der die von der/den Kamera(s) aufgenommenen Daten aufnimmt und verarbeitet und Steuersignale an das Zoom-Objektiv sendet,
- eine mechanische Steuereinrichtung (4), die auf Impulse des Steuerrechners (3) hin den Aufnahmeausschnitt aus dem zu überwachenden Bereich (1) oder die Größe des Bildausschnittes (5) verändern kann, sowie
- eine Einrichtung zum Öffnen einer verschlossenen Passage, die auf Signale des Steuerrechners (3) reagiert.

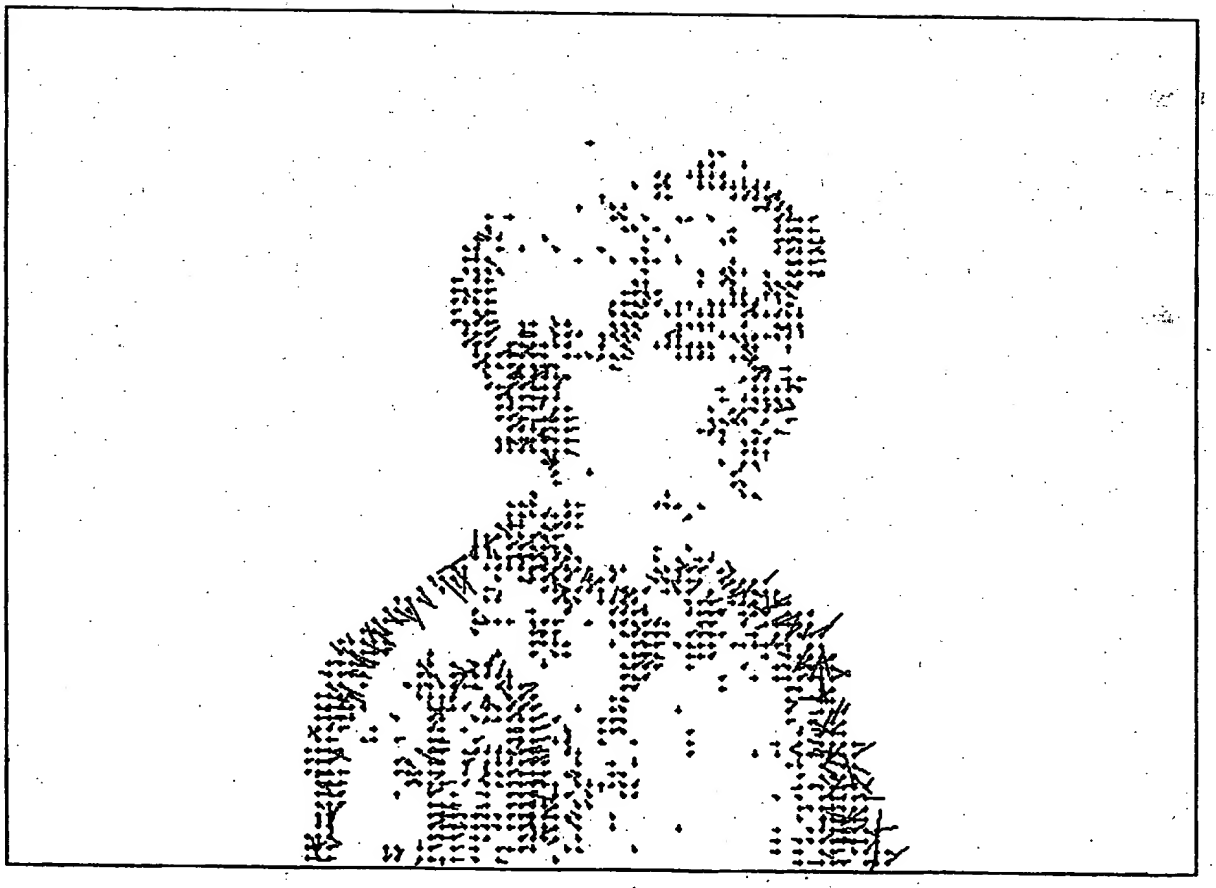
15. Überwachungssystem nach Anspruch 12, zusätzlich umfassend eine Einrichtung zum Erzeugen und Senden von akustischen, optischen und/oder elektrischen (Alarm-) Signalen.

16. Überwachungssystem nach Anspruch 14 oder 15, bei dem zwei Kameras (2) mit ihrem Ruheachse (100) etwa rechtwinklig zu einem Zentrum des räumlichen Bereichs (1) angeordnet sind.

Bild t

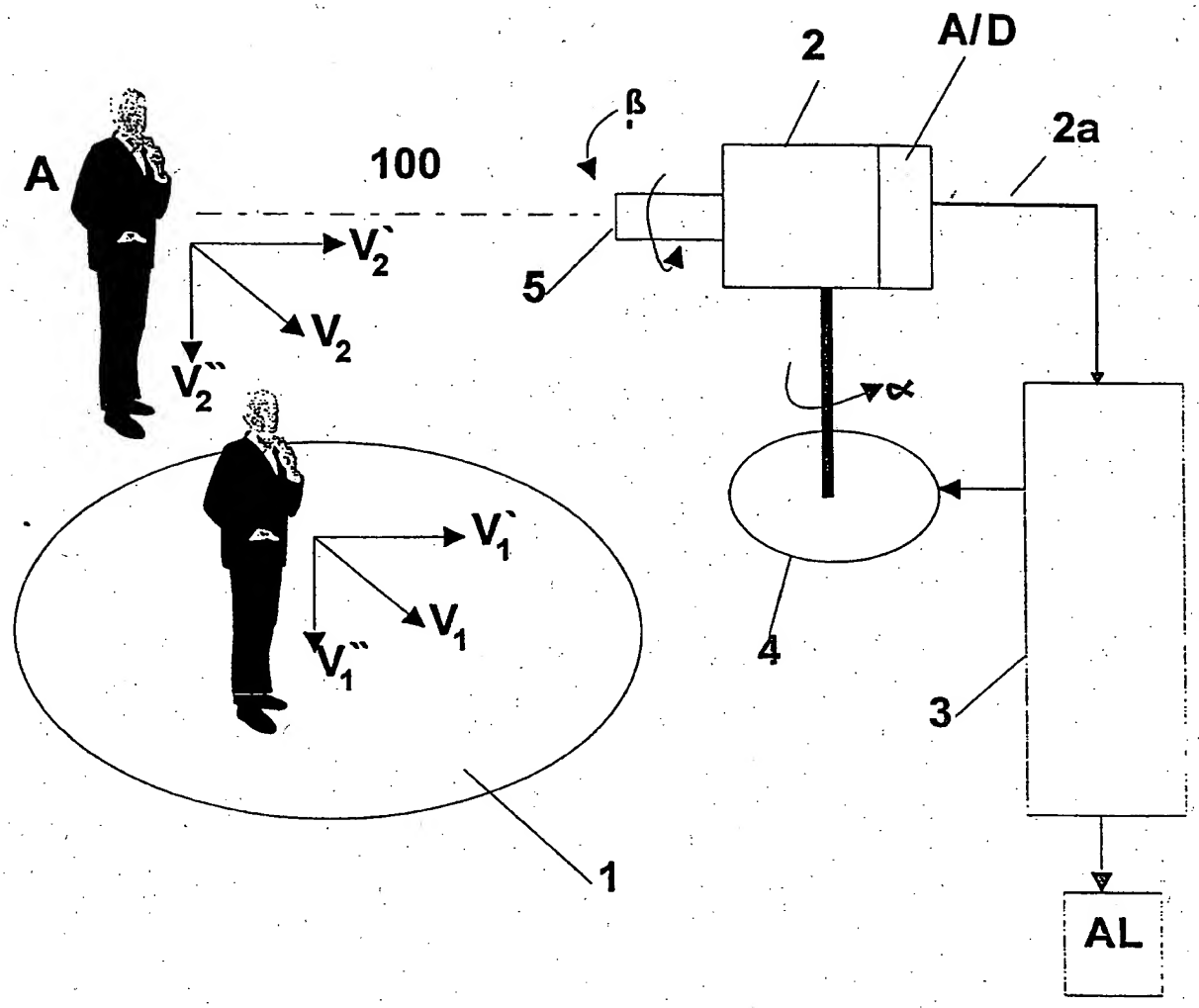
Bild t+1

Figur 1a



Figur 1b

Figur 1c



Figur 2

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/DE 96/01344

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G08B13/196

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 G08B G06T H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PATENT ABSTRACTS OF JAPAN vol. 18, no. 564 (P-1819), 27 October 1994 & JP,A,06 203163 (NEC), 22 July 1994, see abstract ---	1
Y	PROCEEDINGS OF THE ANNUAL INTERNATIONAL CARNAHAN CONFERENCE ON SECURITY TECHNOLOGY, ALBUQUERQUE, OCT. 12 - 14, 1994, no. CONF. 28, 12 October 1994, SANSON L D, pages 31-33, XP000492098 TAKEYUKI TAKANO ET AL: "INTRUDER DETECTION SYSTEM BY IMAGE PROCESSING" see page 32, column 1; figure 1 --- -/--	1

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 November 1996	29. 11. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Breusing, J

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/DE 96/01344

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US,A,5 265 172 (V. MARKANDEY ET AL) 23 November 1993 see column 1, line 21 - column 2, line 37; figure 1	1
A	EP,A,0 445 334 (SIEMENS) 11 September 1991 cited in the application see abstract; figure 3	1
A	DE,A,41 13 992 (W. AMELING ET AL) 5 November 1992 see page 2, line 54 - page 3, line 8; figure 3	11
A	EP,A,0 577 491 (MATRA CAP SYSTÈMES) 5 January 1994 see page 6, line 21 - line 45; figure 4	11
A	EP,A,0 578 558 (THOMSON-CSF) 12 January 1994 see page 2, line 1 - line 24; figure 1	7
A	PATENT ABSTRACTS OF JAPAN vol. 95, no. 8, 29 September 1995 & JP,A,07 134767 (TOSHIBA), 23 May 1995, see abstract	
A	PATENT ABSTRACTS OF JAPAN vol. 18, no. 247 (P-1735), 11 May 1994 & JP,A,06 028474 (FUJITSU), 4 February 1994, see abstract	

1

INTERNATIONAL SEARCH REPORT

Intern. al Application No
PCT/DE 96/01344

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A-5265172	23-11-93	NONE	
EP-A-445334	11-09-91	NONE	
DE-A-4113992	05-11-92	NONE	
EP-A-577491	05-01-94	FR-A- 2693011	31-12-93
EP-A-578558	12-01-94	FR-A- 2693615	14-01-94

INTERNATIONALER RECHERCHENBERICHT

Interna ales Aktenzeichen
PCT/DE 96/01344

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 6 G08B13/196

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 6 G08B G06T H04N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	PATENT ABSTRACTS OF JAPAN vol. 18, no. 564 (P-1819), 27.Oktober 1994 & JP,A,06 203163 (NEC), 22.Juli 1994, siehe Zusammenfassung ---	1
Y	PROCEEDINGS OF THE ANNUAL INTERNATIONAL CARNAHAN CONFERENCE ON SECURITY TECHNOLOGY, ALBUQUERQUE, OCT. 12 - 14, 1994, Nr. CONF. 28, 12.Oktober 1994, SANSON L D, Seiten 31-33, XP000492098 TAKEYUKI TAKANO ET AL: "INTRUDER DETECTION SYSTEM BY IMAGE PROCESSING" siehe Seite 32, Spalte 1; Abbildung 1 --- -/--	1

[X] Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

[X] Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :
"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden
"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist
"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche	Absendedatum des internationalen Recherchenberichts
7.November 1996	29. 11. 96

Name und Postanschrift der Internationale Recherchenbehörde	Bevollmächtigter Bediensteter
Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Breusing, J

Formblatt PCT/ISA/210 (Blatt 2) (Juli 1992)

INTERNATIONALER RECHERCHENBERICHT

Interna ...les Aktenzeichen
PCT/DE 96/01344

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	US,A,5 265 172 (V. MARKANDEY ET AL) 23.November 1993 siehe Spalte 1, Zeile 21 - Spalte 2, Zeile 37; Abbildung 1	1
A	EP,A,0 445 334 (SIEMENS) 11.September 1991 in der Anmeldung erwähnt siehe Zusammenfassung; Abbildung 3	1
A	DE,A,41 13 992 (W. AMELING ET AL) 5.November 1992 siehe Seite 2, Zeile 54 - Seite 3, Zeile 8; Abbildung 3	11
A	EP,A,0 577 491 (MATRA CAP SYSTÈMES) 5.Januar 1994 siehe Seite 6, Zeile 21 - Zeile 45; Abbildung 4	11
A	EP,A,0 578 558 (THOMSON-CSF) 12.Januar 1994 siehe Seite 2, Zeile 1 - Zeile 24; Abbildung 1	7
A	PATENT ABSTRACTS OF JAPAN vol. 95, no. 8, 29.September 1995 & JP,A,07 134767 (TOSHIBA), 23.Mai 1995, siehe Zusammenfassung	
A	PATENT ABSTRACTS OF JAPAN vol. 18, no. 247 (P-1735), 11.Mai 1994 & JP,A,06 028474 (FUJITSU), 4.Februar 1994, siehe Zusammenfassung	

1

INTERNATIONALER RECHERCHENBERICHT

Interna ales Aktenzeichen
PCT/DE 96/01344

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US-A-5265172	23-11-93	KEINE	
EP-A-445334	11-09-91	KEINE	
DE-A-4113992	05-11-92	KEINE	
EP-A-577491	05-01-94	FR-A- 2693011	31-12-93
EP-A-578558	12-01-94	FR-A- 2693615	14-01-94

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 749 098 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
18.12.1996 Bulletin 1996/51

(51) Int. Cl.[6]: **G06T 7/00**

(21) Application number: **96303303.0**

(22) Date of filing: **13.05.1996**

(84) Designated Contracting States:
DE FR GB

(30) Priority: **17.05.1995 JP 118629/95**
07.09.1995 JP 230301/95

(71) Applicant: **HITACHI DENSHI KABUSHIKI KAISHA**
Chiyoda-ku, Tokyo 101 (JP)

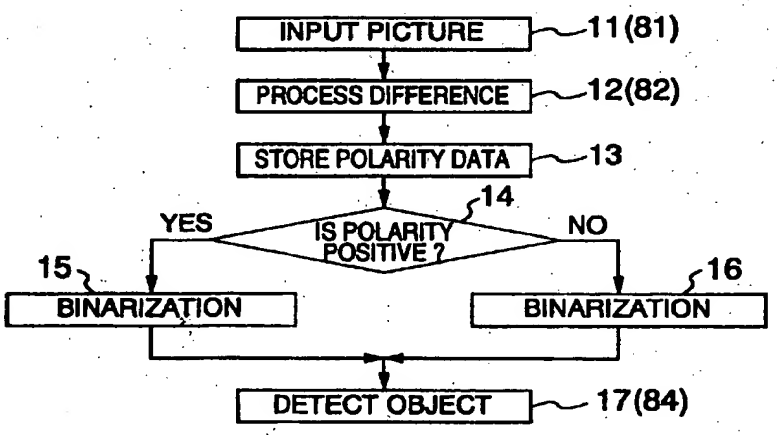
(72) Inventors:
- **Ito, Wataru**
  Kodaira-shi (JP)
- **Ueda, Hirotada**
  Kokubunji-shi (JP)

(74) Representative: **Calderbank, Thomas Roger et al**
MEWBURN ELLIS
York House
23 Kingsway
London WC2B 6HP (GB)

(54) Method and apparatus for sensing object located within visual field of imaging device

(57) An apparatus and a method is provided for sensing an object within an imaging field of an imaging device (121). The data of the reference background picture taken within the imaging field is stored in a storage unit (124). The data of the picture taken by the imaging device (121) is read. The data of the reference background picture is compared with the read picture data at each pixel so that the difference value between both of the data may be yielded. The difference value is compared with a predetermined threshold value at each pixel. Based on the compared result, binarization data about an image of the object is produced. Based on the binarization data, the object is sensed. Then, the data of the reference background picture stored in the storage unit (124) is properly updated with the new data.

FIG.1



EP 0 749 098 A2

Description

The present invention relates to a system for monitoring a picture which system uses an imaging device such as a television camera. More particularly, the invention relates to a method and an apparatus for sensing an object in the system for monitoring a picture which method and apparatus are preferable in the case of needing so high reliability such as the case of sensing an object coming into an dangerous zone.

In recent days, a picture monitoring system arranged to use a TV camera has been made public. Of such monitoring systems, there has been requested a system for automatically sensing an object such as a person or a car coming into a field of a monitor based on a video signal independently of human eyes and issuing a predetermined report and alarm. This type of automatic monitoring system is required to sense an object by processing a video signal produced by a TV camera.

As one of the systems for sensing an object by processing such a video signal, there has been known and widely used as a difference method wherein the latest picture coming to the system in sequence is compared with the one previous picture, a difference of a luminance for each corresponding pixels is obtained, and an area with a large difference as an object area.

The basic flow of the system for sensing a difference value will be described with reference to the flowchart of Fig. 8. As shown, at a step 81, a current picture is captured from a TV camera. At the next step 82, a difference of a luminance is obtained between a pre-stored background picture and the captured picture. Then, at a step 83, binarization process is made for the difference value. At last, at a step 84, the digital value is compared with a predetermined threshold value for sensing if an object exists based on the compared result.

For example, as shown in Fig. 9A, assume that the background picture taken by a TV camera is a railroad crossing. The picture shown in the left hand of Fig. 9B is taken when one car 90 is coming into this crossing. The picture shown in the right hand of Fig. 9B is derived from the binarization of the difference value of the pixel data for each pixel between the background picture of Fig. 9A and the picture shown in the left hand of Fig. 9B.

In this type of system, partial breaks or gaps 91 appear in an object to be sensed as indicated in the picture shown in the right hand of Fig. 9B. These breaks or gaps may bring about an error in sensing an object.

However, this drawback may be overcome by applying a dilation and erosion operation to the system.

As an applied example of a system for sensing a moving object by processing a difference, for example, a writing titled as "Intelligent Image Handling with Image Recognition Technique" written by Ueda, et. al., has been published in "O plus E" No. 176, pp. 122 to 136.

The system for sensing a moving object, published in this writing, is arranged to sense an object on three pictures continuing in time. As shown in Fig. 11, at first, assume that the three continuous pictures 111, 112, 113 in time are input. Luminance differences between the pictures 111 and 112 and between the pictures 112 and 113 are calculated. These two luminance difference values are made binarized and then are subject to the dilation and erosion operation, as a result of which pictures 114 and 115 are obtained. Next, these binary images (pictures) 114 and 115 are ANDed for deriving common portions between these pictures, on which portions a picture 116 of the object is obtained.

In the aforementioned disclosed method, however, the following erroneous sensing may take place. That is, though an object is not actually coming into the view field, an object is erroneously sensed. For example, as shown in Fig. 9C, if the change of a weather or lighting conditions leads to change of brightness of an overall object to be imaged, an object 93 (railroad) may be erroneously sensed though no object is coming into the view field. As shown in Fig. 9D, if an object located out of the view field is made so luminous that a reflected ray 94 may take place in the area within an imaging field, the reflected ray is erroneously sensed as the object 95. Further, as shown in Fig. 9E, if the shadow 96 of the object located out of the field is falling on the area within the imaging field, the shadow may be erroneously sensed as the object 97.

By the way, this differential calculus needs a reference background picture (for example, Fig. 9A). Hence, the reference background picture is created in advance and then pre-stored in a memory. To create the reference background picture, there has been known a system arranged to use a video signal (normally, a video signal of one frame) taken when no object is coming into the view field as the reference background picture according to the operator's selection or a system arranged to average a predetermined number of frames composing a picture at each pixel and to use the averaged frames as the reference background picture.

In the former system, the instantaneous picture is selected on the human determination. If any noise is mingled into the video signal by chance, the noise is held in the background picture. It is thus difficult to obtain the exact background picture. Further, from a view point of the noise circumstances in the transmission system, the mingling of the noise is not so rare. In the practical use, therefore, it is a great obstacle.

In the latter system, to obtain the exact background picture, it is necessary to make the number of frames composing the picture to be averaged as numerous as 1500 frames, for example. As a result, a large time gap (about 50 seconds in a frame frequency of 30 Hz) takes place between a time when a picture for creating the reference background picture is input and a time when a difference is processed for sensing the object. Hence, for example, if the imaging field is gradually made gloomier at dusk, for example, the time gap makes it impossible to create so exact a reference back-

ground picture as allowing the object to be sensed.

It is an object of the present invention to provide a method and an apparatus for reliably and exactly sensing an object coming into a view field of an imaging device for a short time.

It is a further object of the present invention to provide a method and an apparatus for sensing an object which method and apparatus bring about no erroneous sensing resulting from the adverse effect of disturbance factors such as change of a weather or lighting conditions, reflected light or shadow.

It is a still further object of the present invention to provide a method and an apparatus for sensing an object which method and apparatus serve to positively distinguish any portion where a local temperature is changed within the view field of the imaging device from the other portions.

It is another object of the present invention to provide a method and an apparatus for sensing an object with highly reliance by providing a method for automatically creating a suitable reference background picture to deriving a difference value for a short time.

In the method and apparatus for sensing an object coming into the imaging field of an imaging device according to the present invention, the process is executed for storing in a storage unit data of the reference background picture located within the imaging field, reading the data of the picture taken by the imaging device, comparing the data of the reference background picture with the read picture data at each pixel, outputting a difference value of each pixel, comparing the difference value with a predetermined threshold value at each pixel, creating digital data of an object image based on the compared result, sensing if the object exists based on the binary data, and properly updating the data of the reference background picture stored in the storage unit.

The other objects and advantages of the present invention will be apparent from the description that will follow with reference to the appended drawings.

In the drawings

Fig. 1 is a flowchart for describing an operation of an embodiment of the present invention;
Fig. 2 is a flowchart for describing an operation of a further embodiment of the present invention;
Fig. 3 is a flowchart for describing an operation of a still further embodiment of the present invention;
Fig. 4 is a flowchart for describing an operation of another embodiment of the present invention;
Fig. 5 is a flowchart for describing an operation of another embodiment of the present invention;
Fig. 6 is a flowchart for describing an operation of another embodiment of the present invention;
Fig. 7 is a flowchart for describing an operation of another embodiment of the present invention;
Fig. 8 is a flowchart for describing an operation of sensing an object by using a differential calculus;
Figs. 9A, 9B, 9C, 9D, and 9E are pictures for describing a problem caused in sensing an object by using the differential calculus;
Fig. 10 is a view for describing the effect of a similarity used in sensing an object according to the present invention;
Fig. 11 is an explanatory view for sensing an object based on a difference in time according to the prior art;
Fig. 12 is a block diagram showing a system for monitoring a picture to which an embodiment of the present invention is applied;
Fig. 13 is a flowchart showing an embodiment in which the present invention is applied to sensing an object coming into a railroad crossing;
Fig. 14 is a flowchart showing a process for creating a background picture according to an embodiment of the present invention;
Fig. 15 is an explanatory view showing an example of an operation of sensing an object according to an embodiment of the present invention;
Fig. 16 is an explanatory view showing an example of an operation of sensing an object according to an embodiment of the present invention;
Fig. 17 is a flowchart showing an embodiment in which the present invention is applied to a system for monitoring a pyroclastic flow;
Fig. 18 is an explanatory view showing an example of a sensing operation executed by the embodiment in which the present invention is applied to a system for monitoring a pyroclastic flow;
Fig. 19 is a flowchart for describing an operation of creating and updating a reference background picture according to a first embodiment of the present invention;
Figs. 20A and 20B are diagrams for describing an effect caused by calculation of a median of pixel luminance values;
Fig. 21 is a flowchart for describing an operation of creating and updating a reference background picture according to a second embodiment of the present invention;
Fig. 22 is a flowchart for describing an operation of a method for sensing an object according to an embodiment of the present invention;
Fig. 23 is a flowchart showing an operation of an embodiment in which the present invention is applied to sensing an object coming into a railroad crossing;

Figs. 24A, 24B, 24C and 24D are explanatory views showing an operation of sensing an object through the effect of the differential calculus according to an embodiment of the present invention;

Fig. 25 is a flowchart for describing an operation of sensing an object according to another embodiment of the present invention;

Fig. 26 is a flowchart for describing an operation of sensing an object according to yet another embodiment of the present invention; and

Fig. 27 is an explanatory view showing an operation of sensing an object according to the embodiment of the present invention.

Later, the description will be oriented to an embodiment of the present invention with reference to the drawings.

Figs. 1 to 7 are flowcharts showing the operations of the embodiments arranged to sense an object intruding into the inside of a field of an imaging device. The methods according to these embodiments can be executed by the imaging device and a computer and under the control of a program configured to code a shown flowchart stored in a storage medium with a help of the computer.

Before describing the methods shown in Figs. 1 to 7, the description will be oriented to a system for monitoring a picture according to an embodiment of the present invention with reference to Fig. 12.

Fig. 12 shows the picture monitoring system according to the embodiment of the invention. As shown, a TV camera 121 captures a picture of an area to be monitored (monitoring field) and produces a video signal. The video signal is sent to a picture memory 124 through an input interface (I/F) 123 and a data bus 122.

On the other hand, a CPU 126 operates to analyze the picture by using a working memory 125 in accordance with a program stored in a program memory 127 and then starts to sense an object. Based on the process of sensing an object, the program is executed to turn on an alarm lamp 130 through an output interface I/F 128 or display the picture on a monitor 131 through a picture output interface I/F 129. The program for executing the operations of the flowcharts of Figs. 1 to 7 is saved in the program memory 127.

Along the flow shown in Fig. 1, at a step 11 of inputting a picture, a picture is read from a TV camera. Then, at a step 12 of processing a difference, the operation is executed to calculate a difference of each pixel between the read picture and a pre-created reference background picture saved in a memory. Next, at a step 13 of saving code data, a polarity of a difference value of a luminance is saved in a memory. At a next determining step 14, the operation is branched on the polarity. If the polarity is positive, the first binarization process (step 15) is executed. If it is negative, the second binarization process (step 16) is executed. The binarization process is executed to compare the difference value with a threshold value.

These first and second binarization processes (steps 15 and 16) have the corresponding variable threshold values. Hence, if the quality of the change of a luminance against the reference background picture is variable depending on if the polarity is positive or negative, the exact binarization processes can be executed. In accordance with the binarized result, at a step 17, the object is allowed to be positively sensed.

For example, assume that the change of the input picture against the reference background picture is brought about by the change of the weather from cloudy to fine or applying a ray of light to the object. This change causes the polarity of the difference to be positive. Or, if the change is brought about by the shadow cast by the object so that the picture itself is made gloomier, this change causes the polarity of the difference to be negative. The variation depends on the polarity of the difference, that is, if it is positive or negative. Hence, since the prior art operates to simply binarize the absolute value of the difference with a unique threshold value, it does not cope with the wide variation.

However, the method shown in Fig. 1 is arranged to change the direction of the change, that is, the threshold value depending on if the polarity of the difference is positive or negative. Hence, this method enables to flexibly cope with the variation, so that the object may be exactly sensed.

In the flow shown in Fig. 2, a process at a step 24 is added to the process shown in Fig. 1. The process at the step 24 is executed to calculate a similarity of a local luminance distribution pattern between the input picture and the reference background picture. The flow of Fig. 2 is arranged to contain the similarity as the binarization conditions in the first and the second binarization processes (steps 26 and 27). Hence, this processing flow can more flexibly cope with the variation of the input picture against the reference background picture, so that more exact sensing is made possible. The description about the operations of steps 21 to 23, 25 and 28 is left off, because they correspond to the steps 11 to 13, 14 and 17 of Fig. 1, respectively.

Herein, the similarity represents a correlation of the input picture to the reference background picture about the pattern of a local luminance distribution. In Fig. 10, assume that the whole input picture is made gloomier as compared to the reference background picture 101 as indicated by a numeral 102 by the change of illuminance in the imaging environment. In this case, patterns of the background picture 101 such as white lines and railroad are left in the input picture 102 as they are. Hence, a high similarity takes place between a local area of the background picture 101, for example, an area 104 and a local area 105 of the input picture 102.

On the other hand, unlike the reference background picture, as indicated by 103, on the area of the input picture where an object (vehicle) 106 is located, the patterns such as white lines and railroad are disappeared, because the

object 106 hides the background. Hence, a low similarity takes place between the areas 104 and 106.

In the flow of Fig. 3, at the next step 35 to the binarization step 34, the similarity is calculated. At a step 36, the calculated similarity is compared with the reference value. Depending on whether or not the similarity is higher than the reference value, two areas A and B having the different qualities from each other, that is, the area where the luminance of the background picture is changed and the area where the object is located, are separated from each other.

The calculation of the similarity at step 35 is executed only for the pixels at which the difference value exceeds the predetermined threshold value in the binarization process at the step 34. This makes it possible to reduce an amount of calculation to a minimum.

In the flow of Fig. 4, for sensing an object, the process of estimating the similarity is added to the process shown in Fig. 2. If the polarity of the difference is positive, the binarization process is executed at a step 45. Then, at a step 46, the similarity of the form of a local luminance distribution is calculated between the input picture and the reference background picture. At a step 47, the similarity is compared with the reference value. As shown in Fig. 9D, if there exists such an area 94 as being brighter by light applied from a lamp, it is determined that the polarity of the difference is positive and the similarity is high. Hence, at a step 48, it is not determined that any object exists in the area 94. Conversely, if the polarity of the difference is positive and the similarity is not so high (103 of Fig. 10), at a step 49, it is determined that there exists an object in the area 94.

That is, as indicated by the area 94 of Fig. 9D, in the area that is made brighter by light, though the luminance is made higher, the patterns on the road are left. Hence, like the local areas 104 and 105 of Fig. 10, the similarity is made higher.

The use of this similarity, therefore, makes it possible to reliably distinguish the object from the area 94 that is made brighter by a ray of light applied from the lamp.

Next, in the flow of Fig. 5, like Fig. 4, for sensing an object, the process of estimating the similarity is added to the process shown in Fig. 2. If the polarity of the difference is negative, at a step 55, the binarization process is executed. At a step 56, the process is executed to calculate a similarity of the pattern of a local luminance distribution between the input picture and the reference background picture. At a step 57, the similarity is compared with the reference value. As shown in Fig. 9E, if there exists an area 96 that is made dark by the shadow of the object, it is determined that the polarity of the difference is negative and the similarity is high. Hence, at a step 58, it is not determined that any object exists in the area 96 (shadow). Conversely, if the polarity of the difference is negative and the similarity is not so high, at a step 59, it is determined that there exists an object in the area 96.

That is, in the area that is made dark by the shadow of the object itself, as indicated by the area 96 of Fig. 9E, though the luminance is low, the patterns on the road are left. Like the local areas 104 and 105 in Fig. 10, hence, the similarity is made higher.

The use of this similarity, therefore, makes it possible to reliably distinguish the object from the area 96 that is made dark by the shadow.

In the flow of Fig. 6, there is introduced a process (step 67) of estimating a similarity of the pattern of a local luminance distribution between the input picture and the reference background picture. This added process makes it possible to sense an area where a temperature locally goes up.

At a step 61, an object is imaged by an infrared TV camera. The captured video signal is saved as the input picture in a memory. At a step 62, at each pixel, the read input picture is compared with the reference background picture imaged by the same infrared TV camera and pre-stored in the memory for sensing the difference value between both of the pictures and its polarity. Then, at a step 63, the polarity data is stored in a memory. If, at a step 64, the polarity is determined to be positive, at a step 65, the difference value calculated at the step 62 is compared with a predetermined reference value (threshold value) so that the binarization process may be executed. At a step 66, the operation is executed to calculate the similarity between the reference background picture and the input picture. If, at a step 67, the similarity is determined to be low, at a step 68, it is determined that there exists in the input picture an area (object) where a temperature goes up.

That is, a difference of a luminance distribution takes place between the area where the local temperature goes up and the area where it keeps stable. The similarity of the form of the luminance distribution between the input picture and the reference background picture is made low. For example, assume that the method for sensing an object as shown in Fig. 6 is applied to sensing a pyroclastic flow. If a lava dome is formed by the pyroclastic flow, the temperature abruptly goes up in the local area of the lava dome. Hence, the similarity is made lower. On the other hand, for example, as to the clouds that are made warmer by the pyroclastic flow, the temperature goes up uniformly for the whole clouds. Hence, the similarity is made higher.

Hence, the process (step 67) of estimating the similarity is added to the process of Fig. 1 in order that the warmed clouds may be distinguished from the pyroclastic flow.

In the flow of Fig. 7, there is introduced a process (step 77) of estimating a similarity of the form of a local luminance between the input picture and the reference background picture. This process makes it possible to sense a local area where the temperature goes down.

At a step 71, an object is imaged by an infrared TV camera. The video signal captured by the TV camera is read

as an input picture in a memory. At a step 72, the reference background picture taken by the infrared TV camera and pre-stored in a memory is compared with the read input picture at each pixel so that the difference value between both of the pictures and its polarity may be sensed. At a step 73, the polarity data is stored in a memory. If the polarity is determined to be negative at a step 74, at a next step 75, the difference value calculated at the step 72 is compared with the predetermined reference value (threshold value) so that the binarization process may be executed. At a step 76, the process is executed to calculate a similarity between the reference background picture and the input picture. If, at a step 77, the similarity is determined to be low, at a step 78, it is determined that there exists an area (object) whose temperature goes down in the input picture.

That is, a difference of a luminance distribution takes place between the local area whose temperature goes down and the local area whose temperature is inhibited to go down. The similarity of the pattern of a luminance distribution is made lower between the input picture and the reference background picture. For example, consider that the method for sensing an object as shown in Fig. 7 is applied to sensing a pyroclastic flow. If the lava dome is fell down, the temperature of the lava dome is abruptly made lower, because the heat source of the lava dome is dropped. It means that the similarity is lower. On the other hand, for example, as to the clouds warmed by the pyroclastic flow, if the application of heat to the clouds by the heat source is reduced, the whole clouds are substantially uniformly cooled down. Similarly, the temperature of the clouds may uniformly go down. Hence, the similarity is made higher.

At a step 77 of estimating the similarity, therefore, the process makes it possible to separate only the object (area) whose temperature locally goes down from the other objects in sensing the objects.

Later, the method for sensing an object according to the present invention will be described in detail along the illustrative embodiments.

The below-described embodiments are arranged so that the present invention is applied to a picture monitoring system for sensing an object such as a walker or a car coming into a railroad crossing. A TV camera is installed at a site where the camera can take a look at the whole railroad crossing. The TV camera produces a video signal representing the picture taken by the camera itself. The video signal is processed for sensing an object. The description about the embodiment will be expanded along the flowcharts. The operation of the flowchart about the embodiment is executed by the system shown in Fig. 12.

Fig. 13 shows an object sensing process executed by the CPU 126 according to an embodiment of the invention. The process executed by this embodiment is a combination of the processes shown in Figs. 4 and 5 to which a process of updating a background picture is added.

When the process of Fig. 13 is started, at a step 141 of inputting a picture, a video signal picked by a camera 121 is read as the image data f(x, y) consisting of 320 x 240 pixels, where x and y are coordinates of pixels on the screen and f(x, y) represents a luminance value of a pixel at the coordinates (x, y).

Next, at a determining step 142, it is determined whether or not the new reference background picture is created.

In order to create and update this reference background picture, the CPU 126 executes the process shown in Fig. 14 in parallel to the flow of Fig. 13.

Entering the flow of Fig. 14, at first, at a step 161, the input picture are captured from the TV camera at regular intervals of 0.2 second. When the frames of the captured picture reach 20, those 20 frames are saved (stored) in a memory.

Next, at a step 162, at each pixel, the median of each of the 20 frames is calculated by the following expression (1) for creating the background picture data r(x, y).

$$r(x,y) = \underset{1 \le i \le 20}{med}\{f^{(i)}(x,y)\} \qquad \cdots\cdots (1)$$

wherein $f^{(i)}(x, y)$ represents the picture data for 20 frames and i = 1, 2, ..., 20.

In the expression (1), med { } represents a median. The pixels at the same location of the 20 frames are ranged in greater order of their luminance values. Taking the pixel of the tenth (middle) value, the pixel is assumed as the background picture data r(x, y). At a step 163, this expression (1) is calculated for deriving the new reference background picture.

Hence, the reference background picture 163 is made equivalent to the picture data r(x, Y).

In this embodiment, only 20 frames of the picture are sampled at regular intervals of 0.2 seconds. The frame period of the imaging device such as the normal TV camera is 1/60 second. To read the picture on which the median calculation is done, of 60 frames per one second output from the TV camera, any number of frames may be read at proper intervals if necessary. That is, the output frames of the imaging device are properly thinned out for reading them. Or, the picture may be read at a longer period than the output period of the frame of the imaging device. We will further explain this matter later.

Going back to the process of Fig. 13, if the creation of the background picture is terminated, at a step 143, the existing reference background picture stored in the memory is updated to a new reference background picture that has been

just created. Then, the operation goes to the next step 144. If the creation of the new reference background picture is not terminated, the operation goes to the step 144 of processing a difference. At a step 144, the input picture is divided into blocks each consisting of 2 x 2 pixels. The overall picture is divided into a block area $f_n(i, j)$ consisting of 160 x 120 pixels. Then, an average luminance Af of each block area is derived, where n is a block number and i, j represents a coordinate location of a pixel inside of the block.

Further, for the reference background picture data r(x, y), likewise, the overall picture is divided into a block area $r_n(i, j)$ consisting of 160 x 120 pixels. Then, an average luminance Ar of each block area is derived. The average luminances A_f and A_r of each block area are calculated by the following expression (2);

$$A_f = \frac{1}{[B]} \sum_{i,j \in B} f_n(i,j) \qquad (2)$$

$$A_r = \frac{1}{[B]} \sum_{i,j \in B} r_n(i,j)$$

where B represents the subject block area, [B] represents a pixel number contained in the area B. (This embodiment uses a block consisting of 2 x 2 pixels. Hence, i, j∈B represents $0 \leq i \leq 1$ and $0 \leq j \leq 1$ and [B] is four.)

Next, a difference value ε_n of an average luminance between the input picture and the reference background picture is derived at each block by the expression (3).

$$\varepsilon_n = |Af - Ar| \qquad (3)$$

$$(n = 1, 2, 3 \ldots)$$

At the succeeding step 145 of saving the polarity data, the polarity data S_n of the difference is derived by the expression (4) and then is saved in a memory.

$$S_n = \begin{cases} 1 & (Af > Ar) \\ -1 & (Af \leq Ar) \end{cases} \quad \cdots\cdots \qquad \cdots\cdots (4)$$

Next, at a determining step 146, the operation proceeds to the first or the second binarization step 147 or 152 according to the polarity data S_n.

These binarization processes are executed by comparing the difference value ε_n with a predetermined threshold value. The threshold value for the positive polarity S_n, that is, the threshold value th+ at the step 147 is derived by the expression (5).

$$th^+ = \mu^+ + K^+ \sigma^+ \qquad (5)$$

, where

$$\mu^+ = \frac{\sum_{d \in R^+} \varepsilon_d}{[R^+]} \qquad \sigma^+ = \sqrt{\frac{\sum_{d \in R^+} (\varepsilon_d - \mu^+)^2}{[R^+]}}$$

where ε_d represents a difference value between the input pixel and the background picture at the d-th block (see the expression (2)),

μ^+ is an average difference value of the blocks having positive difference values of the blocks predetermined to contain no object, this value representing an average level of noise contained in the picture,

σ^+ is a standard deviation of the difference values of the blocks having positive difference values of the blocks predetermined to contain no object, this value representing how variable the level of the noise contained in the picture is,

and

K+ is a adjustable parameter and [R+] is a number of pixels contained in an area wherein the polarity Sn is positive and no object exists.

Further, at the step 152, likewise, the threshold value th^- for the negative polarity Sn is derived by the following expression (6).

$$th^- = \mu^- + K^- \sigma^- \tag{6}$$

, where

$$\mu^- = \frac{\sum_{d \in R^-} \varepsilon_d}{[R^-]} \quad \sigma^- = \sqrt{\frac{\sum_{d \in R^-} (\varepsilon_d - \mu^-)^2}{[R^-]}}$$

where K^- is an adjustable parameter and $[R^-]$ is a number of pixels contained in an area having a negative polarity and no object.

At the steps 148 and 153 of calculating the similarity, for each block, the similarity γ_n is calculated by the following expression (7).

$$\gamma_n = \frac{-\sum_{d \in D_n} \sum_{ij \in B} |(f_d(i,j) - f^-_d) - (r_d(i,j) - r^-_d)|}{[B] \cdot [D_n]} \tag{7}$$

where D_n represents an area near an n-th block and [ ] represents a number of elements.

At steps 149 and 154 of determining the similarity, the adjustable parameters T^+ and T^- are set as the threshold values for these steps. On these parameters, it is determined whether or not the similarity γ_n is higher than the threshold value.

Concretely, at the determining step 149, the similarity γ_n is compared with the threshold value T^+. If the similarity γ_n exceeds the threshold value T^+, at the step 150, it is determined that the area that is lit by the lamp is sensed as indicated in Fig. 9D as the area 94 from which the light is reflected. Only if the similarity γ_n is smaller than or equal to the threshold value T^+ at the step 149, it is determined that the object is sensed at the step 151.

Likewise, at the determining step 154, the similarity γ_n is compared with the threshold value T^-. If γ_n exceeds the threshold value T^-, at the step 155, it is determined that the area that is made dark by the shadow of the object is sensed as indicated in Fig. 9E as the shadow area 96. Only if γ_n is smaller than or equal to the threshold value T^-, at the step 156, it is determined that the object is sensed.

Then, when the object is sensed at the step 151 or 156, the CPU 126 operates to turn on the alarm lamp 130 through the output I/F 128 or display the picture on the monitor 131 through the picture output I/F 129 for reporting to the user that the object is sensed.

As mentioned above, the method according to this embodiment is capable of positively sensing only an object within the field of the imaging device. The concrete arrangement to which the method for sensing an object as shown in Fig. 13 will be described with reference to Figs. 15 and 16.

In Fig. 15, on an input picture 171, at dusk, a car 171A is stopped near a railroad crossing. The shadow 171B of the car 171A extends to the inside of the crossing. If this input picture 171 is determined by the conventional method, as indicated on a picture 173, it is erroneously determined that there exists an object 173A coming into the crossing.

On the other hand, according to the embodiment shown in Fig. 13, at first, the process shown in Fig. 14 is executed to obtain a reference background picture 172 and update it to a new background picture every about four seconds.

Next, since the portion of the shadow 171B makes the code data S_n negative, the operation goes to the determining step 154 at which the similarity between the reference background picture 172 and the input picture 171 is determined.

In the area where the car 171A exists, since the car 171A hides a portion of the background, the similarity is not made so high. At the step 156 shown in Fig. 13, therefore, this area 171A is determined to actually have an object. On the contrary, in the area of the shadow 171B, nothing hides the background. Hence, the similarity is made high. As a result, at the step 155 of Fig. 13, it is determined that it is not an object but the shadow.

According to this embodiment, therefore, as shown by the picture 174, the car 174A is clearly distinguished from the shadow 174B, so that only an object coming into the crossing is positively sensed. In this case, though the shadow

174B appears, it is determined that no object comes into the crossing.

Next, Fig. 16 shows an input picture 181 where a ray of light 181B applied by a head lamp of a motorcycle 181A is intruded into the railroad crossing before the motorcycle enters into the crossing. In this case, if the input picture 181 is determined by the conventional method, as shown by the picture 183, it is erroneously determined that there exists an object 183A coming into the crossing.

On the other hand, according to the embodiment shown in Fig. 13, at first, the process shown in Fig. 14 is executed to obtain a reference background picture 172 and update it to a new picture every about four seconds.

Next, since the polarity data S_n is positive in the ray of light 181B applied by the head lamp, the operation goes to the step 149 of determining the similarity. At this step 149, it is determined whether or not the similarity between the latest reference background picture 172 and the input picture 181 is high.

In the area where the motorcycle 181A exists, since the motorcycle 181A hides a portion of the background, the similarity is not made so high. Hence, at the step 151, it is determined that there exists an object. On the other hand, in the area of the ray 181B, since it does not hide any portion of the background, the similarity is made high. As a result, at the step 150, it is determined that it is not an object but a mere bright portion.

According to this embodiment, therefore, as shown by the picture 184, the motorcycle 174A and the ray applied from its head lamp are clearly distinguished from each other as the objects 151 and 150 shown in Fig. 13. It means that only the object coming into the railroad crossing can be positively sensed. In this case, though the portion of the ray 184B appears, it is determined that there exists no object coming into the railroad crossing.

Next, the description will be oriented to a method according to another embodiment of the present invention.

The arrangement of this embodiment is application of the present invention to monitoring a pyroclastic flow in an active volcano. In the apparatus for monitoring a picture as shown in Fig. 12, an infrared TV camera (infrared imaging device) is used for the camera 121. The infrared TV camera yields the input picture on which a luminance is made larger in a warmer portion of an object. In the system shown in Fig. 12, the CPU 126 executes the process shown in Fig. 17.

In Fig. 17, as depicted in brackets, the routines from steps 191 to 199 and from steps 201 to 203 are the same as those from the steps 141 to 149 and from the steps 152 to 154 shown in Fig. 13. Hence, these routines will be briefly described. The processing operation goes to any one of the step 199 of determining the similarity and the step 203 of the same function, depending on the luminance, that is, the polarity of the difference of the temperature between the input picture and the reference background picture. At the steps 199 and 203, the adjustable parameters T^+ and T^- are used as the threshold values for determining the similarity. If an area has a positive polarity and a similarity that is smaller than or equal to the threshold value, at the step 200, it is determined that the area (object) has a higher temperature. If the area has a negative polarity and a similarity that is smaller than or equal to the threshold value, at the step 204, it is determined that the area (object) has a lower temperature.

The operation of the method according to this embodiment will be described with reference to Fig. 18. An input picture 211 is a picture captured at an instance when a pyroclastic flow takes place. A numeral 212 denotes a reference background picture.

Then, after the process of Fig. 17 is executed for the picture shown in Fig. 18, the resulting picture is made to be the picture 213. On this picture, a portion whose temperature is made higher by the pyroclastic flow is allowed to be positively distinguished from the portion whose temperature is made lower after the pyroclastic flow is flown out.

That is, since the difference of the luminance distribution pattern takes place between the area whose local temperature is higher and the other areas whose temperatures keep stable, the similarity of the form of the luminance distribution is made lower.

In the portion where a lava dome is formed by the pyroclastic flow, the local temperature abruptly goes up. Hence, the similarity is made lower. On the other hand, for example, in the clouds that are made warmer by the pyroclastic flow, the whole temperature uniformly goes up. Hence, the similarity is made higher.

Hence, the addition of the step 199 of estimating the similarity makes it possible to clearly distinguish the area (pyroclastic flow) whose temperature goes up from the other area at the step 200 if there exist portions such as the heated clouds.

On the other hand, in the area whose local temperature goes down, the difference of the luminance distribution pattern takes place between that area and the area whose temperature keeps stable. Hence, the similarity of the pattern of the luminance distribution is made lower.

If the lava dome is fell down, it means that the heat source is dropped. The temperature of the relevant area is made abruptly lower. Hence, the area is made to have a lower similarity. On the other hand, in the area such as the clouds heated by the pyroclastic flow, if the heat applied by the heat source is reduced, the overall area is substantially uniformly cooled down, that is, lowers its temperature. Hence, the similarity is made higher.

The addition of the step 203 of estimating the similarity, therefore, makes it possible to clearly distinguish the area whose temperature is made lower (pyroclastic flow is flown out) from the other area at the step 204 if there exist portions such as the clouds whose temperature is made lower. It means that the cooled area can be distinctly sensed.

In addition, it goes without saying that the present invention is not limited to the above-mentioned application but may be applicable to another system such as a system for monitoring an intruder to a prohibitive zone such as plant

facilities.

According to the embodiment as described above, the sign of the difference value is used for overcoming the conventional shortcoming of the erroneous sensing resulting from the change of the background caused by the variation of illuminance. This thus makes it possible to lower the adverse effect caused by the change of the environment such as change of illuminance of the monitoring field, thereby being able to accurately sense an object. Further, the similarity of the pattern of the luminance distribution is used for determining how much of degree the patterns are saved. Hence, if a phantom area caused by the change of the illuminance takes place though no actual object exists, such an area is allowed to be positively distinguished from an actual object when sensing an object.

As noted above, the method according to the present invention may be reliably applied to such an area as requiring high reliance like a railroad crossing. It means that this method may make great contribution to widening the applied field of the apparatus for monitoring a picture.

In the embodiment as described above, for generating the reference background picture, the method shown in Fig. 14 as well as another method may be used. The reference background picture may be produced for calculating a median of plural picture frames as well as calculating an average value of plural frames though the latter method needs more frames. Moreover, another statistical method may be used therefor. A method for generating a reference background picture by calculating a median will be described below with reference to the drawings.

The method according to the embodiment shown in Fig. 19 includes the step 301 of receiving a picture, a step 302 of storing the picture, a step 303 of determining whether or not a predetermined number of frames have been saved, a step 304 of calculating a median for each pixel of the picture saved when the storage for the predetermined number of frames is terminated, and a step 305 of updating the picture calculated at the step 304 to a new reference background picture.

In Fig. 19, at the step 301, the picture is read from the TV camera at a rate of 30 frames per second or a higher sampling rate. Then, the read picture is stored in the memory at the step 302.

Afterwards, if, at the step 303, the number of the stored frames reaches a predetermined number N (N is a positive integer), at the step 304, a median value of a luminance for each pixel is calculated on the picture signal for the stored N frames. Then, at the step 305, the picture composed to have the median value as the luminance of each pixel is updated to a new reference background picture and then saved in the memory.

As this result, the method of this embodiment is capable of positively securing a quite exact reference background picture without having to set a large value as the frame number N (for example, N = 300). The ground will be described below.

Remarking a pixel (x, y) of the picture whose N frames are stored, the N frames are ranged in numerical order. The change of the luminance of the pixel is made to be as shown by 310 in Fig. 20A. Herein, (x, y) represents a pixel location (coordinates) within a frame. It means that the picture data composed of 320 x 240 pixels has the values of x = 1 to 320 and y = 1 to 240.

Next, about the pixel (x, y) at one location, a median value of the luminance is calculated. The sentence "A median value of the luminance is calculated" means that pieces of data ranged in an order are rearranged in order of magnitude from a smaller to a larger (ascending order) or from a larger to a smaller (descending order), so as to select a central value at the middle of the order.

As a result of calculating a median of the luminance of each pixel as shown in Fig. 20A, the distribution as shown by a characteristic 320 in Fig. 20B takes place. The median value of the pixel (x, y) is the N/2-th luminance value M, that is, a median value M.

If no change of the picture takes place over N frames, the luminance value of the same pixel of each frame is made equal to the median value M. In this case, the characteristic 320 is made to be a horizontal line.

Then, assume that an object is caught or any noise is mingled on L frames ($1 < L < N$) of the stored N frames though the object or noise is not on the other frames rather than the L frames. The pixels composing the object or the pixels mingled with the noise is made to have a different luminance value from the pixels of the picture that does not suffer from any change, that is, a different median value M therefrom. By rearranging the luminance values of the pixels at the same location of the N frames in descending order, if the frame number L is equal to or smaller than N/2, the pixel is located at both ends of the characteristic 320 of Fig. 20B spaced from the center location (N/2).

When the luminance values of the pixels composing the object or mingled with noise are smaller than those of these pixels that do not suffer from any change, the pixels are located in the left hand (toward a value of 1) from the center location (N/2). Conversely, when the former is larger than the latter, the pixels are located in the right hand (toward a value of N) from the center location (N/2).

As the number L of the frames where an object or any noise is caught is made greater, the pixels appearing close to the center location (N/2) is made more numerous. While the frame number L is N/2 or less, no substantial influence effect is applied to the median value M.

That is, the influence applied by the object or any noise takes place in the area around the center location (N/2) close to the median value M only when the pixels where the object or any noise is caught stay at the same location over at least a half of a number N of stored frames.

In principle, however, a moving object or noise is quickly moved or disappears. It means that the number of frames where the object or noise is caught is not so great.

Hence, by calculating a median value, the influence of the moving object or noise may be restricted without having to read so many frames N. This calculation of the median value makes it possible to obtain a quite exact reference background picture. In addition, the calculation of the median value just needs a half or less of the stored frames if an object is caught in some of the frames. Hence, by lowering a sampling rate, the frames to be stored may be reduced in number.

Next, the method according to another embodiment of the invention shown in Fig. 21 is arranged to add an object sensing step 402 and a determining step 403 after the step 401 of receiving a picture. After it is determined whether or not an object exists at the steps 402 and 403, the same process as described about the method of the embodiment shown in Fig. 19 is executed only for the picture that is determined to have no object. In addition, the description about the processes at the steps 401 and 404 to 407 is left out, because those processes are the same as those at the steps 301 to 305. At the step 402, by using a differential calculus, it is determined whether or not an object comes into an imaging field. Then, the picture that is determined to have no object at the step 403 is stored in the memory at the step 404. Next, if the number of stored frames reaches the predetermined number, the processes at the step 406 of calculating a median value and the step 407 are executed to calculate a median value (step 406) and update a reference background picture (step 407).

According to this embodiment, therefore, it is guaranteed at the step 403 that a new background picture used for updating the reference background picture at the step 407 does not have any object coming into the picture. This thus makes it possible to greatly reduce the number N of frames required for calculating a median value to a small value such as N = 6 as compared to the value used in the embodiment shown in Fig. 19.

Next, in the method according to another embodiment shown in Fig. 22, the processes done by the embodiments shown in Figs. 19 and 21 are allowed to be executed in parallel. At the steps 305 and 407, that is, two-system processes of updating a reference background picture, the method of this embodiment is capable of constantly yielding the latest reference background picture.

In the embodiment shown in Fig. 19, the number of frames to be stored is about 300, which is far more numerous than that of the embodiment shown in Fig. 21. Hence, the interval for updating the reference background picture is made considerably wide. It means that the freshness of the reference background picture is not necessarily satisfactory.

On the other hand, in the embodiment shown in Fig. 21, if yes at the step 403 of determining whether or not an object exists, the process at the step 407 of updating a reference background picture is not executed. In this case, it is possible to make no use of such a special advantage as reducing the necessary number of frames.

The method according to the embodiment shown in Fig. 22, however, can make use of both advantages of the embodiments shown in Figs. 19 and 21 and allow the processes for these embodiments to use the corresponding latest reference background picture. Hence, the reference background picture can be made more exact.

Fig. 23 shows a series of processes from a process of creating a reference background picture to a process of sensing an object, which processes are executed by the CPU 126 in the system shown in Fig. 12. This embodiment is based on the method shown in Fig. 19.

At a step 501 of reading a picture, a video signal is read as picture data consisting of 320 x 240 pixels from a TV camera 121. Then, at a step 502 of storing a picture, the picture data is stored in a picture memory 124.

At a determining step 503, if the number of the frames being stored reaches a predetermined value N_1 (for example, $N_1 = 300 = N$), the operation is branched into a step 504 of calculating a median. If it does not reach the value N1, the operation is branched into a step 506 of sensing an object.

At the step 504 of calculating a median, the process is executed to rearrange the stored N frames as shown in Fig. 20 for each pixel and calculate a median value M by the following expression (8).

$$r(x,y) = \underset{1 \le i \le 300}{med} \{f^{(i)}(x,y)\} \quad \cdots\cdots (8)$$

where $f^{(i)}(x, y)$ represents the picture data composed of 300 frames and i takes a value of i = 1, 2, ..., 300. In addition, the expression (8) is the same as the case that the number N of the frames in the expression (1) is 300.

At the step 505, the median value derived at the step 504 is stored in a storage area for the reference background picture allocated to the picture memory 124.

Next, at the step 506, the process is executed to derive a difference between the picture data read at the step 501 and the reference background picture stored in the storage area of the picture memory 124. If the picture data has a portion with a larger difference than a predetermined threshold value, the portion is sensed as an object.

At first, the reference background picture as shown in Fig. 24A is compared with the picture containing an object to be sensed as shown in Fig. 24B. At each pixel located at the same coordinates of both of the pictures, the difference of the luminance is derived. The picture composed on the difference value is made to be as shown in Fig. 24C. Then, the difference value is given a predetermined threshold value and then is made binarized. As a result, the picture shown in

Fig. 24D is extracted.

On the other hand, if no object exists, it does not appear as the extracted picture. Hence, only the object is allowed to be sensed.

At a step 507, it is determined whether or not an object exists based on the result given at the step 506. If any object exists, at a step 508, an alarm is issued. If no object exists, the operation is branched into the step 501 of reading a picture.

In turn, Fig. 25 shows an applied one of the method described with reference to Fig. 21. Each figure in brackets of Fig. 25 indicates the step of another Figure at which the same process as the figure-attached step is executed.

In the embodiment shown in Fig. 25, the processes at the steps 601 to 605 the same as those at the steps 501 to 505 in Fig. 23. Hence, the description about these steps is left off this specification.

At the step 606, like the step 601, the video signal is read as the picture data consisting of 320 x 240 pixels from the TV camera 121 to the picture memory 124.

At a step 607 of sensing an object, if a new reference background picture is created at a step 612, the new one is used. If not, a reference background picture created at the step 605 is used for sensing an object.

At a step 608, if it is determined that no object exists at the step 607 of sensing an object, the operation is branched to a step 609, while if it is determined that any object exists, the operation is branched to a step 613.

If it is determined that no object exists, at the step 609, the picture obtained at the step 606 is stored in the memory.

At a determining step 610, if the number of frames being stored reaches a predetermined number N_2 (for example, $N_2 = 6$), the operation is branched to a step 611 of calculating a median value. If the number does not reach the predetermined value N_2, the operation is branched to a step 606 of reading a picture.

At the step 611 of calculating a median value, the operation is executed to rearrange the predetermined number N_2 of the stored frames as shown in Fig. 20 for each pixel and calculate a median value by the above-indicated expression (8).

The median value derived at the step 611 is stored in a storage area for a reference background picture allocated to the picture memory 124 at a step 612.

Fig. 26 shows an applied one of the method described with reference to Fig. 22.

The processes at the steps 701 to 705 are the same as the steps 501 to 505 of Fig. 23. Hence, the description about these processes is left off the specification.

In succession, at a step 706, the process of reading a picture is executed. This is the same as the process of the step 501. And, at a step 707, the process of sensing an object is executed. This is the same as the process of the step 506. Then, a first and a second systems of processes are executed in parallel. The first system of process is the same as the method shown in Fig. 19. The second system of process is the same as the method shown in Fig. 21.

The processes of steps 708 to 710 corresponding to the first system of process are the same as the processes of the steps 502 to 504 shown in Fig. 23. Next, the processes of steps 712 to 715 and 717 corresponding to the second system of process are the same as the processes of the steps 608 to 611 shown in Fig. 25. Hence, the description about those steps is not left off the specification. In step 711, a picture obtained by the first system of the process is taken as a candidate A of the reference background picture. In step 716, a picture obtained by the second system of the process is taken as a candidate B of the reference background picture.

At steps 718, 719 and 720, the processes are executed to select an earlier one of a reference background picture A obtained at the step 711 for the first system of process and a reference background picture B obtained at the step 716 for the second system of process and replace the previous reference background picture stored in the memory with the selected one. A number N_0 of frames read at the step 703, a number N_1 of frames read at the step 709, and a number N_2 of frames read at the step 714 are set to have the relation of $N_0 \geq N_1 \geq N_2$. If the relation of $N_0 > N_1 > N_2$ is set, a proper number of frames to the imaging condition are sampled so that an object may be efficiently sensed. For example, a recommended value of N_0 is $N_0 = 500$.

In turn, the description will be oriented to the sensing results given by those embodiments with reference to Fig. 27.

In Fig. 27, a picture 801 is an example of picture data that is taken by the TV camera 121 (see Fig. 12) and read at the picture reading step 501.

For such a picture, according to the embodiment shown in Fig. 23, unless a truck in the picture stays at one site for about five or more seconds (N = 300 at a frame frequency of 60 Hz), an exact reference background picture where no object (truck) appears can be positively obtained as shown by the picture 802 of Fig. 27.

In the embodiments shown in Figs. 25 and 26, unless the object is erroneously sensed for at least three frames, as shown by the picture 802 of Fig. 27, an exact reference background picture where no object (truck) appears can be obtained.

According to the above-mentioned embodiments, therefore, if the frames required for creating the reference background picture are reduced in number, the exact reference background picture can be positively obtained. Hence, even in the conditions like an outdoor place where an environmental change such as variable illuminance, the latest reference background picture with little delay may be used for sensing an object. This makes it possible to offer a highly reliable monitoring system.

In addition to the above-mentioned applications, the present embodiment may be applicable to a system for monitoring an intruder to plant facilities or prohibitive zones.

According to the present embodiment, the calculation of a median is used for creating the reference background picture. Hence, the method of this embodiment can overcome the drawback the conventional technique of creating the reference background picture has involved and is endurable in light of noise. Further, the method of this embodiment needs a smaller number of frames for creating the reference background picture as compared with the method arranged to take an average of each pixel therefor. Hence, the method of this embodiment offers a highly reliable monitoring system with the latest reference background picture.

In addition, an object sensing method according to the present invention may use luminance data as well as color data of the video signal.

Claims

1. An apparatus for sensing an object within an imaging field of an imaging device, comprising:

   storage means (124) for storing data of a reference background picture fitted in said imaging field;
   means (123) for reading data of a picture taken by said imaging device;
   means (125, 126) for comparing the data of said reference background picture data with said read picture data at each pixel and outputting a difference value of each pixel;
   means (126) for updating the data of said reference background picture saved in said storage means;
   means (126) for comparing said difference value with a predetermined threshold value at each pixel and producing binarization data about said object based on the compared result; and
   means (126) for sensing whether or not said object exists based on said binarization data.

2. An apparatus as claimed in claim 1, wherein said means for outputting said difference value contains means for sensing a polarity of said difference value and said means for producing digital data contains first binarization means having a threshold value for a positive polarity of said difference value and second binarization means having a threshold value for a negative polarity of said difference value.

3. An apparatus as claimed in claim 2, further including means for comparing the data of said reference background picture with said taken picture data at each block consisting of a predetermined number of pixels and calculating a similarity of said data between the blocks of said both of data, and said means for sensing whether or not said object exists contains means for sensing intrusion of said object based on said similarity and said binarization data.

4. An apparatus as claimed in claim 3, wherein the data of said reference background picture and said current taken picture data contain data about luminance of plural pixels composing said picture.

5. An apparatus as claimed in claim 4, wherein said imaging device includes a television camera (121).

6. An apparatus as claimed in claim 4, wherein said imaging device includes an infrared imaging device.

7. An apparatus as claimed in claim 1, wherein said means for reading said data operates to read said picture data having a predetermined number of frames from said imaging device at a predetermined period, said means for updating the data of said reference background picture contains means for calculating a median of said read picture data composed of the predetermined number of frames at each pixel and operates to replace the data of said reference background picture saved in said storage means with said median value of each pixel as the new data of said reference background picture.

8. An apparatus as claimed in claim 7, wherein said predetermined number of frames read from said imaging device by said means for reading the data is smaller than the number of frames output by said imaging device at said predetermined period.

9. An apparatus as claimed in claim 7, wherein the period at which said data reading means reads frame data from said imaging device is longer than the frame period of output of said imaging device.

10. An apparatus as claimed in claim 7, wherein the data of said reference background picture and the data of said taken picture contain data about luminance of plural pixels composing the picture.

11. An apparatus as claimed in claim 7, wherein said means for sensing whether or not said object exists contains

means (130) for issuing an alarm if said object is sensed and said updating means operates to update the data of said reference background picture if said sensing means does not sense said object.

12. An apparatus as claimed in claim 11, wherein said means for reading the data contains first data reading means for reading picture data composed of N1 frames and second data reading means for reading picture data composed of less N2 frames than said N1 if said object is not sensed,

said means for calculating a median contains first median calculating means for calculating a median value of said read picture data composed of the N1 frames at each pixel and second median calculating means for calculating a median value of said read picture data composed of the N2 frames at each pixel, and

said updating means operates to replace the data of said reference background picture saved in said storage means with the median value of any one of said first and second median calculating means as the new data of said reference background picture.

13. A method for sensing an object within an imaging field of an imaging device, comprising the steps of;

updating (143) data of a reference background picture within said imaging field, said data stored in storage means;
reading (141, 142) data of a picture taken by said imaging device;
comparing (12) the data of said reference background picture stored in said storage means with said read picture data at each pixel for calculating a difference value (εn) of each pixel;
comparing (15, 16) said difference value with a predetermined threshold value at each pixel and producing binarization data about the picture of said object based on the compared result; and
sensing (17) whether or not said object exists based on said binarization data.

14. A method as claimed in claim 13, wherein when calculating said difference value, a process is executed to sense a polarity of said difference value, and at the step of producing said digital data, a process is executed to use a first threshold value (th^+) if the polarity is positive (15) or a second threshold value (th^-) if the polarity is negative (16).

15. A method as claimed in claim 14, further including the step of comparing the data of said reference background picture with the data of said taken picture at each block consisting of the predetermined number of pixels and calculating a similarity (γn) of the data value between both of said data blocks based on the compared result, and in which said object is sensed on said similarity and said digital data.

16. A method as claimed in claim 15, wherein said step of calculating the similarity is located after said step of creating the digital data and the similarity is calculated only for the pixels whose difference values depart from said threshold value.

17. A method as claimed in claim 16, further including a step of comparing said calculated similarity with a predetermined reference value and at said step of sensing the object, it is determined that said object exists if said similarity is higher than said reference value.

18. A method as claimed in claim 17, wherein the data of said reference background picture and the data of said taken picture contain data about luminance of plural pixels composing the picture, said object similarity is a value for representing a similarity of distribution pattern of local luminance values of said picture, and at said step of sensing said object, the process is executed to determine as existence of no object an area having pixels whose difference values have a positive polarity and whose similarities are higher than said reference value (48).

19. A method as claimed in claim 17, wherein the data of said reference background picture and the data of said taken picture contain data about luminance of plural pixels composing the picture, the similarity of said object is a value for representing the similarity of a distribution pattern of local luminance values of said picture, and at said step of sensing said object, a process is executed to determine as existence of no object an area having pixels whose difference values have a negative polarity and whose similarities are higher than said reference value (58).

20. A method as claimed in claim 16, further including a step (36) of comparing said calculated similarity with a predetermined reference value and steps (37, 38) of distinguishing an area having pixels whose similarities are higher than said reference value from an area having pixels whose similarities are lower than said reference value.

21. A method as claimed as claim 20, wherein said imaging device is an infrared imaging device, the data of said reference background picture and the data of said taken picture contain data about luminance of plural pixels com-

posing the picture, the similarity of said object is a value for representing the similarity of a distribution pattern of local luminance values of the picture, and at said step of distinguishing said areas from each other, a process is executed to determine as an area having a higher temperature than another area an area having pixels whose difference values have a positive polarity and whose similarities are smaller than or equal to said reference value (68).

22. A method as claimed in claim 20, wherein said imaging device is an infrared one, the data of said reference background picture and the data of said taken picture contain data about luminance of plural pixels composing the picture, the similarity of said object is a value for representing the similarity of a distribution pattern of local pixel values of said picture, and at said step for distinguishing said areas from each other, a process is executed to determine as an area whose temperature is made lower than the other areas an area having pixels whose difference values have a negative polarity and whose similarities are smaller than or equal to said reference value (78).

23. A method as claimed in claim 13, wherein at said step of reading the data, said picture data composed of the predetermined number of frames is read from said imaging device at a predetermined period, and said step of updating the data of said reference background picture contains a step (304) of calculating a median (M) of said read picture data composed of the predetermined number of frames at each pixel and a step (305) of replacing the data of said reference background picture stored in said storage means with said median value of each pixel.

24. A method as claimed in claim 23, wherein at said step of reading the data, said predetermined number of frames read from said imaging device is smaller than the number of frames output by said imaging device at said predetermined period.

25. A method as claimed in claim 23, wherein at said step of reading the data, the period at which the frame data is read from said imaging device is longer than the period of the frame output by said imaging device.

26. A method as claimed in claim 23, wherein said step of sensing the object contains a step of issuing an alarm if said object is sensed and at said updating step, a process is executed to update the data of said reference picture if said object is not sensed.

27. A method as claimed in claim 23, wherein said step of reading the data contains a first data reading step (603) of reading picture data composed of N1 frames and a second data reading step (610) of reading picture data composed of N2 frames that is less than said N1 if said object is not sensed (N1 and N2 are both a positive integer of 2 or more),

said step of calculating the median contains a first median calculating step (604) of calculating a median of said read picture data composed of the N1 frames at each pixel and a second median calculating step (611) of calculating a median of said read picture data composed of the N2 frames at each pixel, and

at said step of updating the data, a process is executed to replace the data of said reference background picture stored in said storage means with a median value derived by any one of said first and second median calculating steps as the data of said reference background picture.

FIG.1

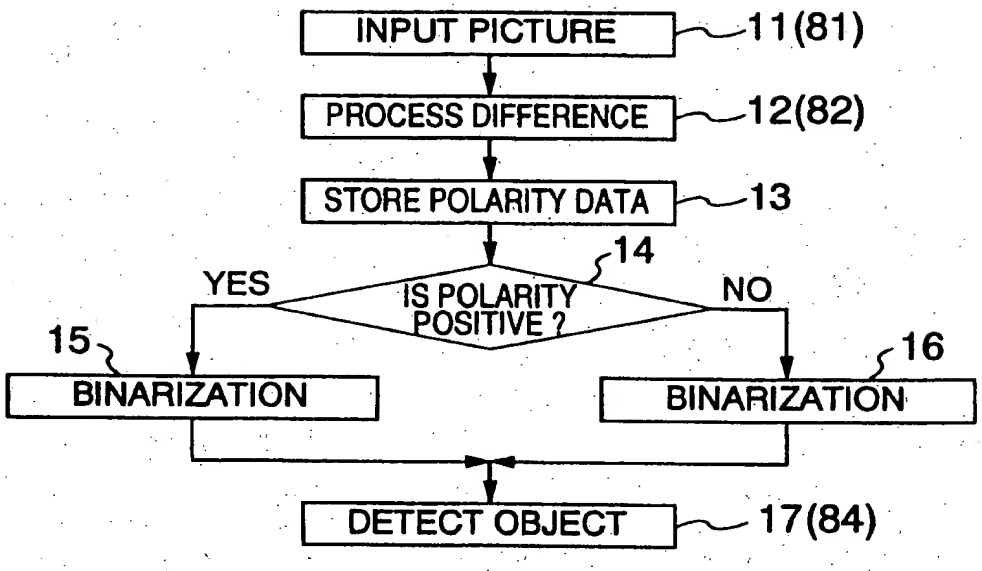

INPUT PICTURE
11(81)
PROCESS DIFFERENCE
12(82)
STORE POLARITY DATA
13
14
IS POLARITY POSITIVE ?
YES
NO
15
BINARIZATION
16
BINARIZATION
DETECT OBJECT
17(84)

FIG.2

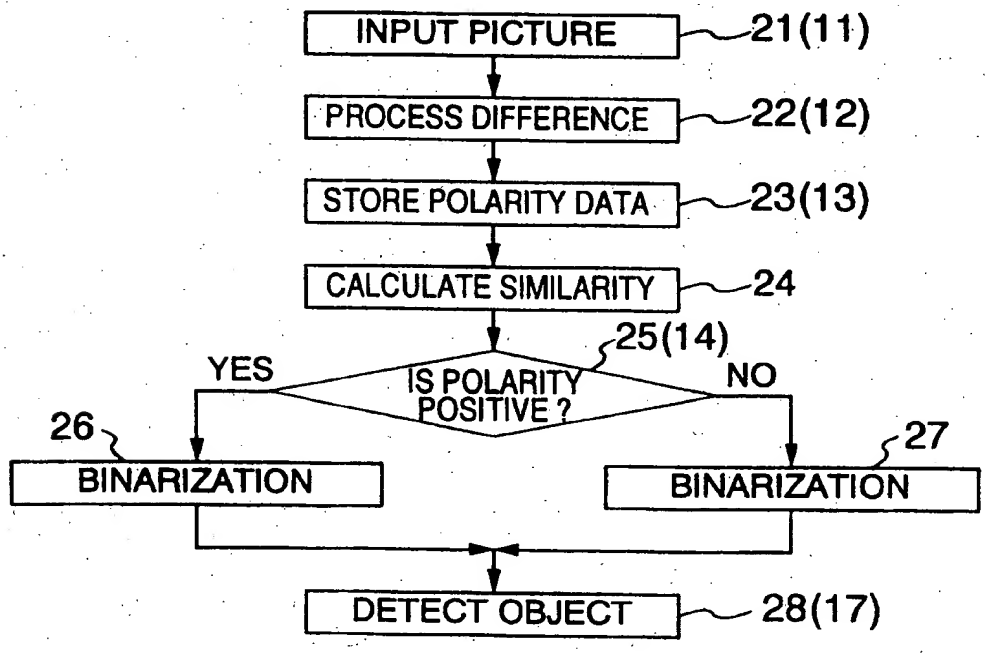

INPUT PICTURE
21(11)
PROCESS DIFFERENCE
22(12)
STORE POLARITY DATA
23(13)
CALCULATE SIMILARITY
24
25(14)
IS POLARITY POSITIVE ?
YES
NO
26
BINARIZATION
27
BINARIZATION
DETECT OBJECT
28(17)

FIG.3

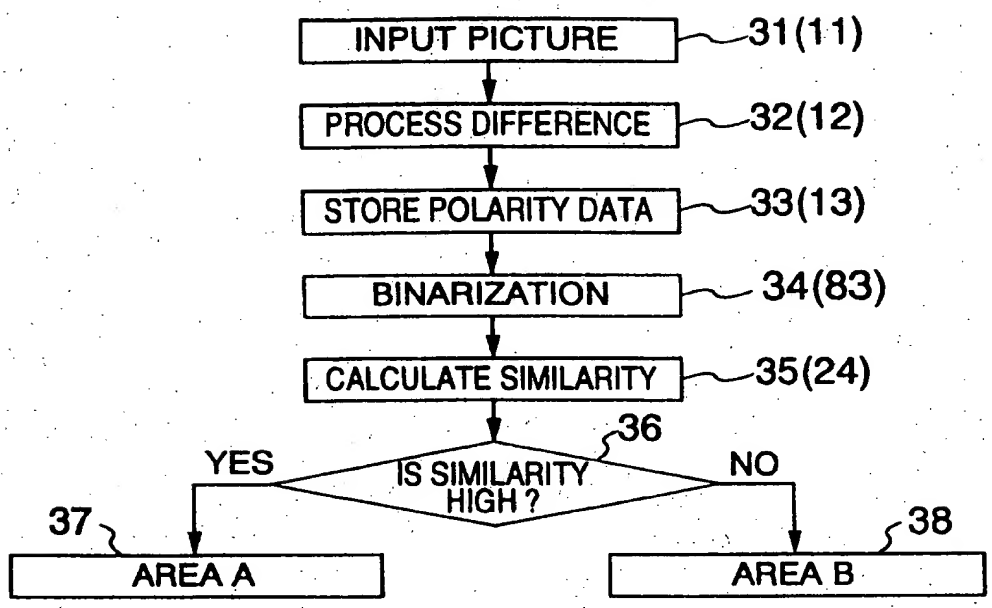

INPUT PICTURE
31(11)
PROCESS DIFFERENCE
32(12)
STORE POLARITY DATA
33(13)
BINARIZATION
34(83)
CALCULATE SIMILARITY
35(24)
36
IS SIMILARITY HIGH ?
YES
NO
37
AREA A
38
AREA B

FIG.4

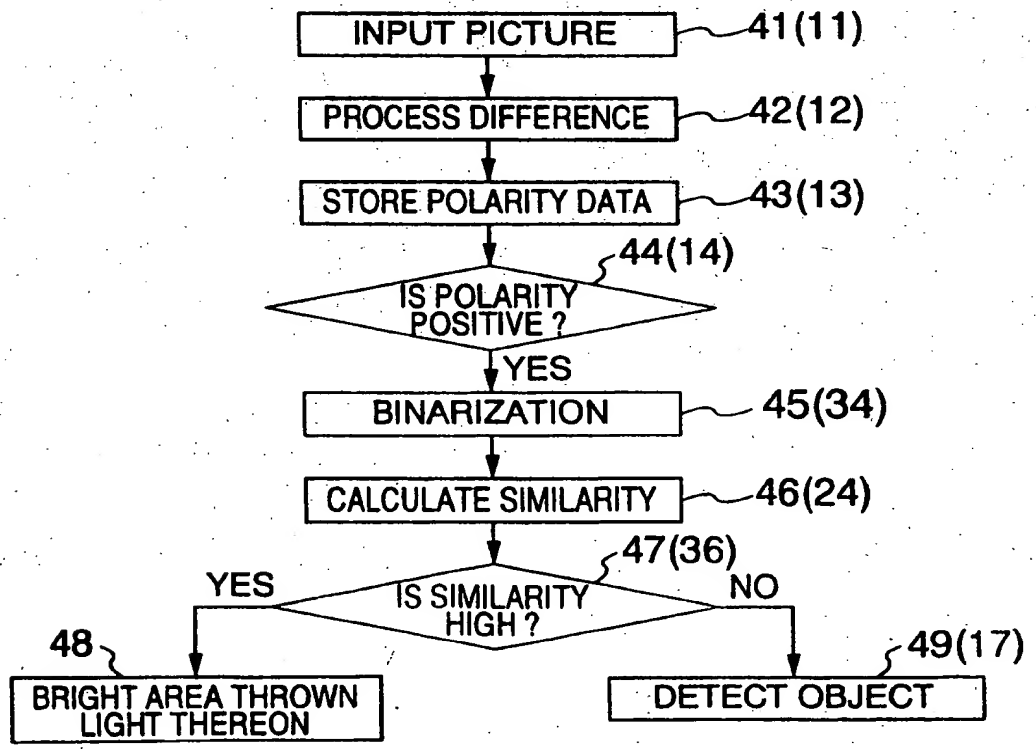

INPUT PICTURE
41(11)
PROCESS DIFFERENCE
42(12)
STORE POLARITY DATA
43(13)
44(14)
IS POLARITY POSITIVE ?
YES
BINARIZATION
45(34)
CALCULATE SIMILARITY
46(24)
47(36)
IS SIMILARITY HIGH ?
YES
NO
48
BRIGHT AREA THROWN LIGHT THEREON
49(17)
DETECT OBJECT

FIG.5

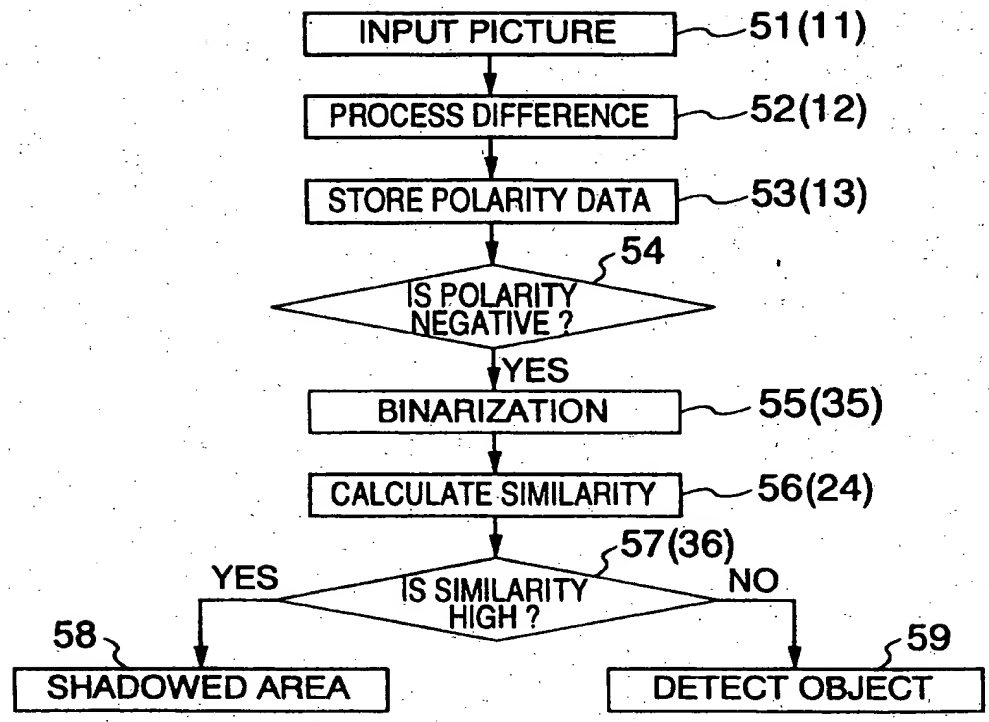

INPUT PICTURE
51(11)
PROCESS DIFFERENCE
52(12)
STORE POLARITY DATA
53(13)
54
IS POLARITY NEGATIVE ?
YES
BINARIZATION
55(35)
CALCULATE SIMILARITY
56(24)
57(36)
IS SIMILARITY HIGH ?
YES
NO
58
SHADOWED AREA
59
DETECT OBJECT

FIG.6

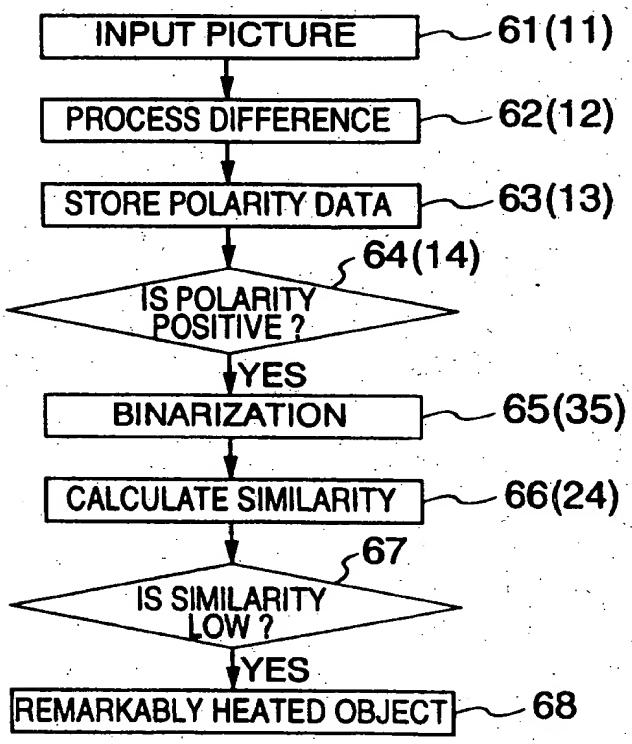

INPUT PICTURE
61(11)
PROCESS DIFFERENCE
62(12)
STORE POLARITY DATA
63(13)
64(14)
IS POLARITY POSITIVE ?
YES
BINARIZATION
65(35)
CALCULATE SIMILARITY
66(24)
67
IS SIMILARITY LOW ?
YES
REMARKABLY HEATED OBJECT
68

FIG.7

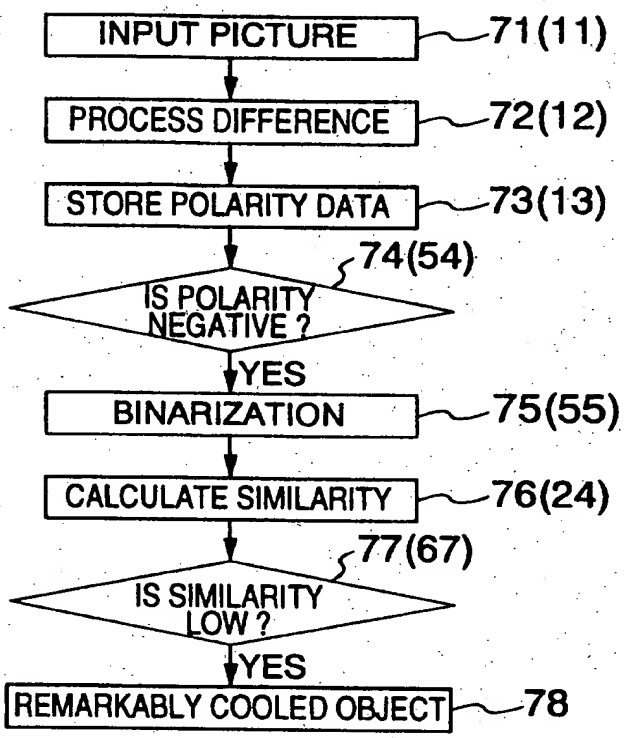

INPUT PICTURE
71(11)
PROCESS DIFFERENCE
72(12)
STORE POLARITY DATA
73(13)
74(54)
IS POLARITY NEGATIVE ?
YES
BINARIZATION
75(55)
CALCULATE SIMILARITY
76(24)
77(67)
IS SIMILARITY LOW ?
YES
REMARKABLY COOLED OBJECT
78

FIG.8

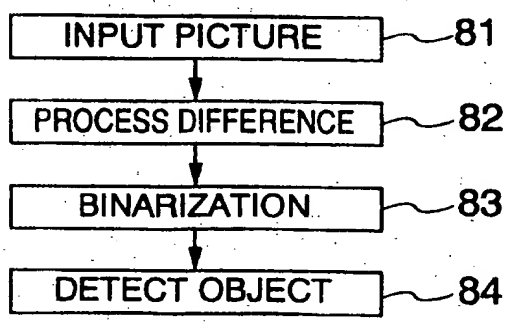

INPUT PICTURE
81
PROCESS DIFFERENCE
82
BINARIZATION
83
DETECT OBJECT
84

FIG.9A

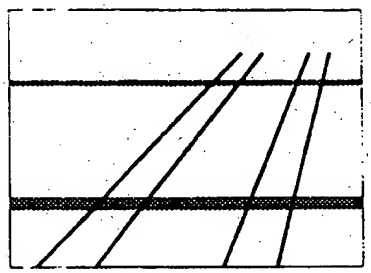


FIG.9B

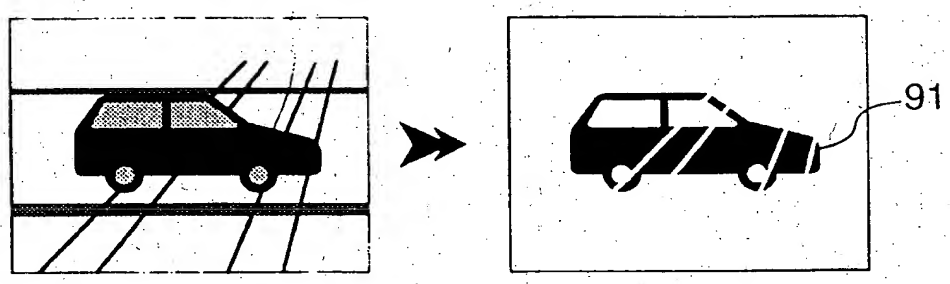


FIG.9C

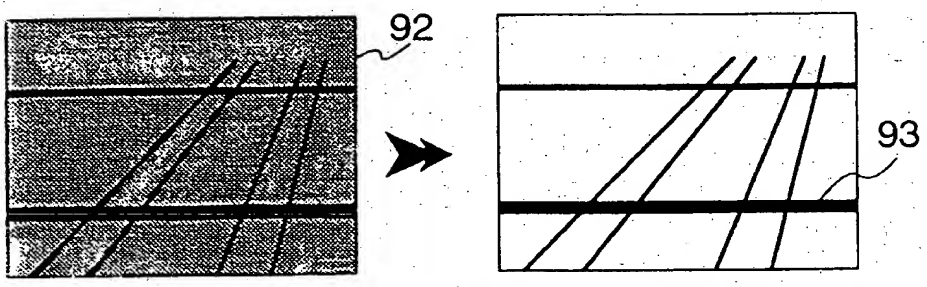


FIG.9D

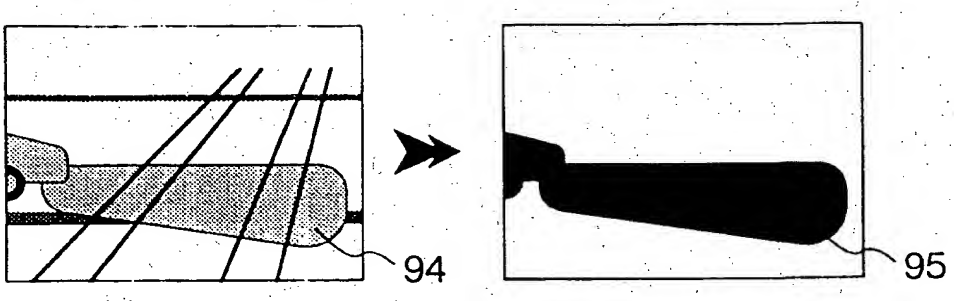


FIG.9E

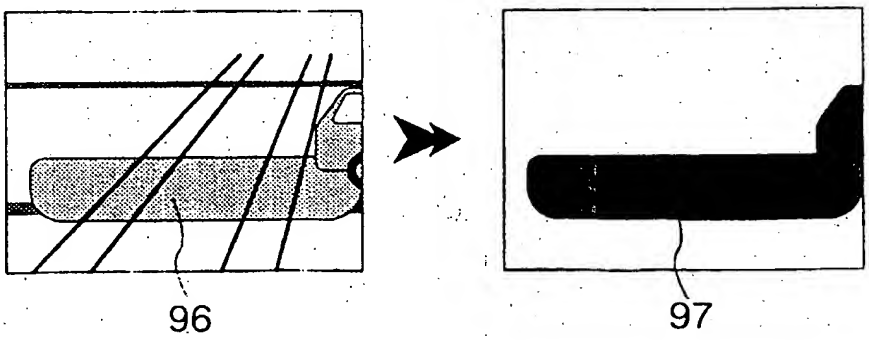


FIG.10

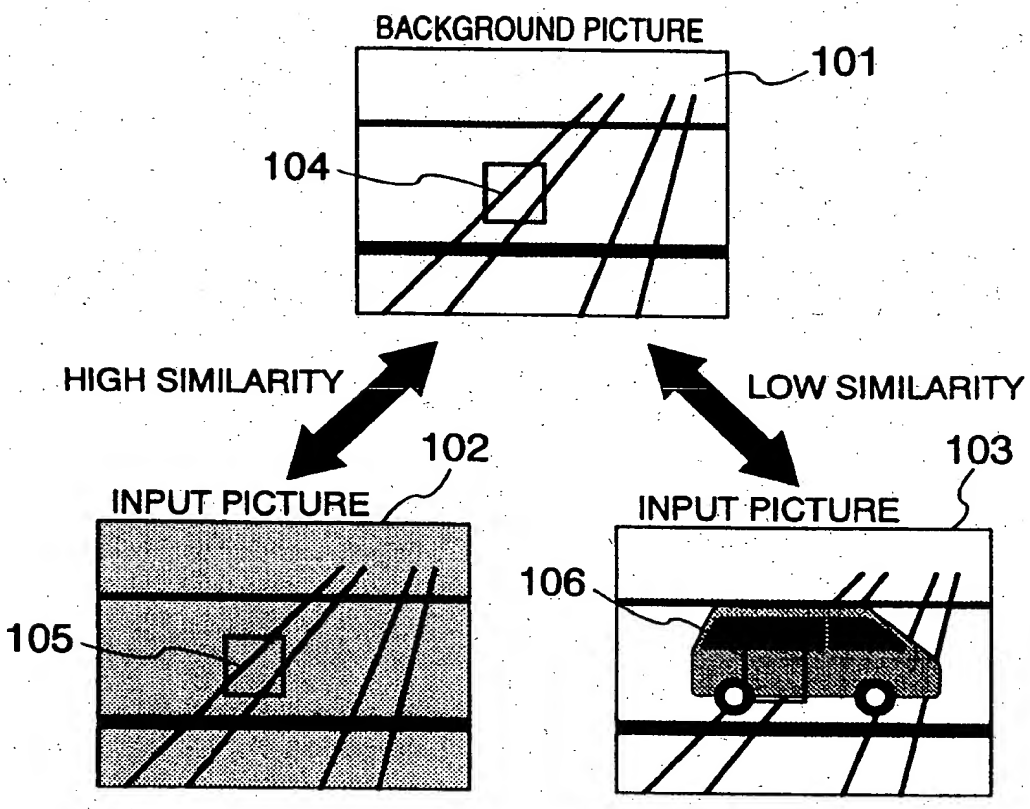

BACKGROUND PICTURE
101
104
HIGH SIMILARITY
LOW SIMILARITY
102
INPUT PICTURE
103
INPUT PICTURE
105
106

FIG.11

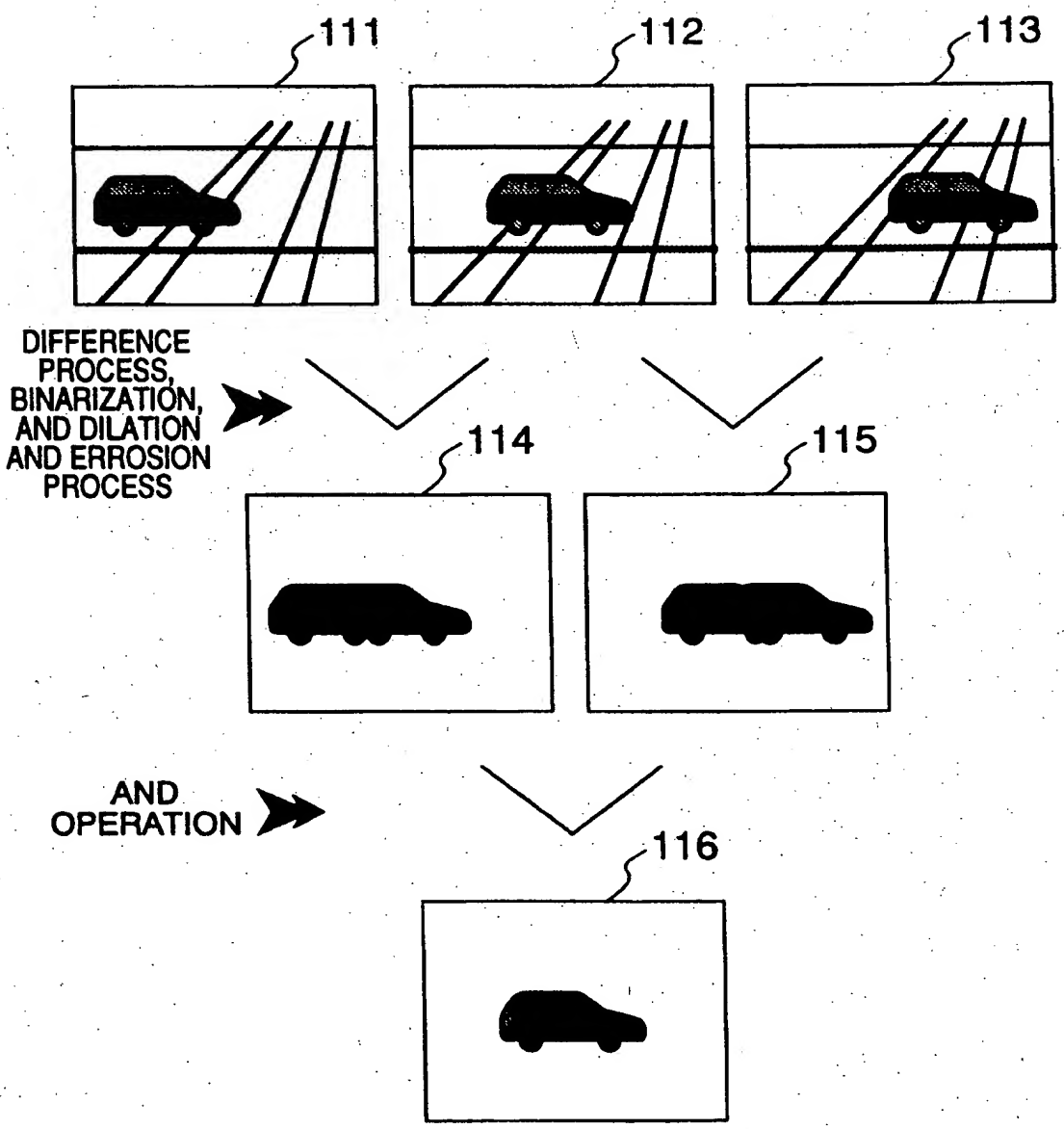

111
112
113
DIFFERENCE PROCESS, BINARIZATION, AND DILATION AND ERROSION PROCESS
114
115
AND OPERATION
116

FIG.12

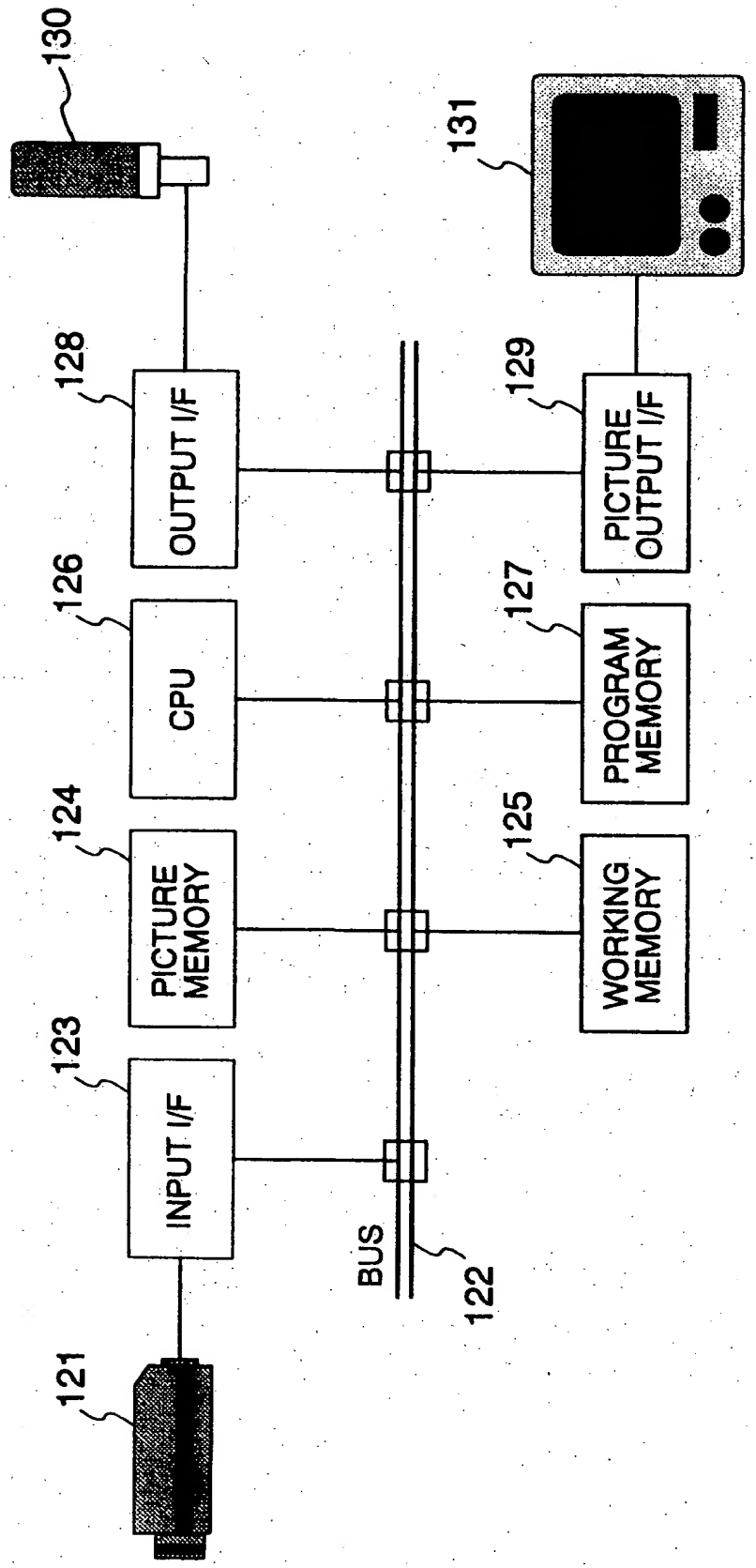

130
131
128
OUTPUT I/F
129
PICTURE OUTPUT I/F
126
CPU
127
PROGRAM MEMORY
124
PICTURE MEMORY
125
WORKING MEMORY
123
INPUT I/F
BUS
122
121

FIG.13

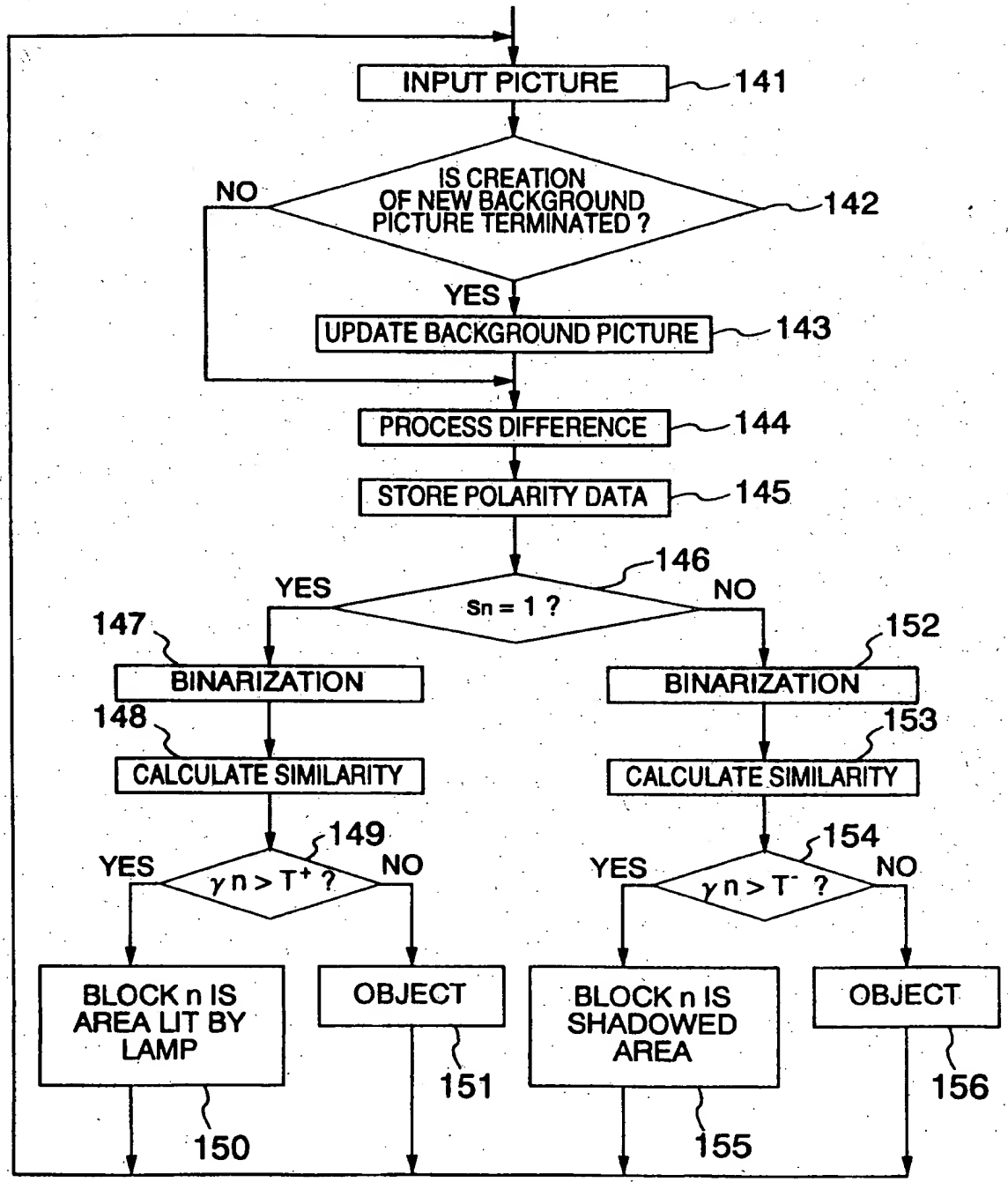

INPUT PICTURE
141
IS CREATION OF NEW BACKGROUND PICTURE TERMINATED ?
142
NO
YES
UPDATE BACKGROUND PICTURE
143
PROCESS DIFFERENCE
144
STORE POLARITY DATA
145
146
Sn = 1 ?
YES
NO
147
BINARIZATION
148
CALCULATE SIMILARITY
149
γn > T+ ?
YES
NO
BLOCK n IS AREA LIT BY LAMP
150
OBJECT
151
152
BINARIZATION
153
CALCULATE SIMILARITY
154
γn > T- ?
YES
NO
BLOCK n IS SHADOWED AREA
155
OBJECT
156

FIG.14

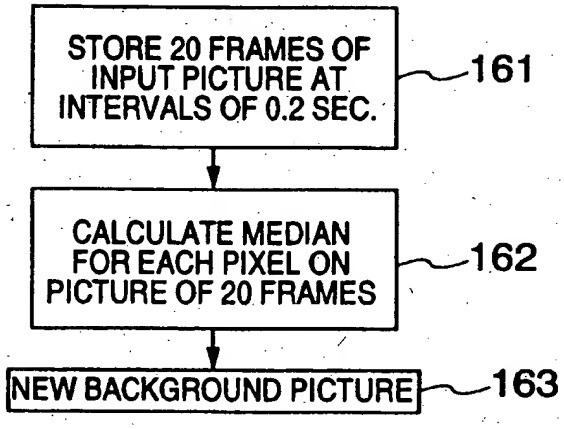

STORE 20 FRAMES OF INPUT PICTURE AT INTERVALS OF 0.2 SEC.
161
CALCULATE MEDIAN FOR EACH PIXEL ON PICTURE OF 20 FRAMES
162
NEW BACKGROUND PICTURE
163

FIG.15

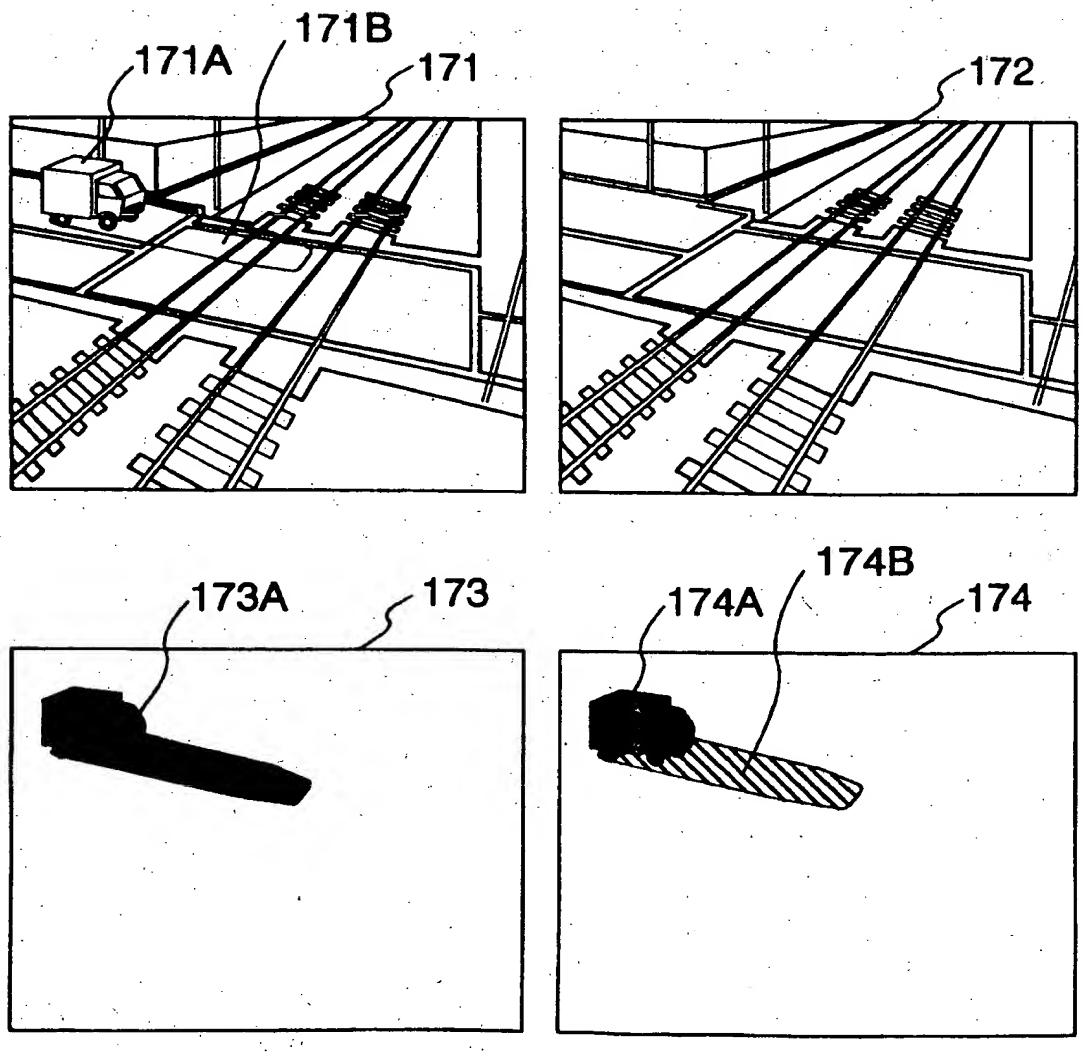

171A
171B
171
172
173A
173
174B
174A
174

FIG.16

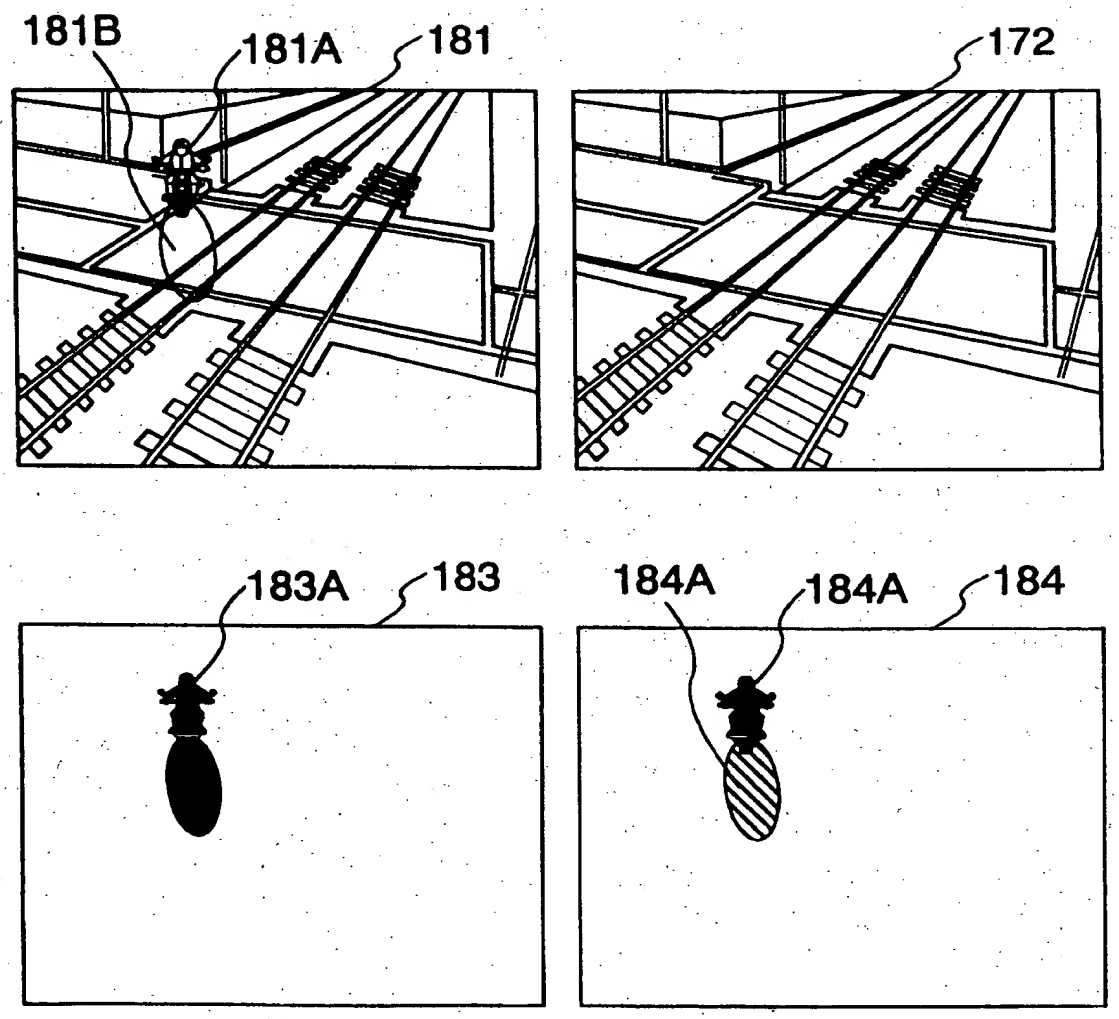


FIG.17

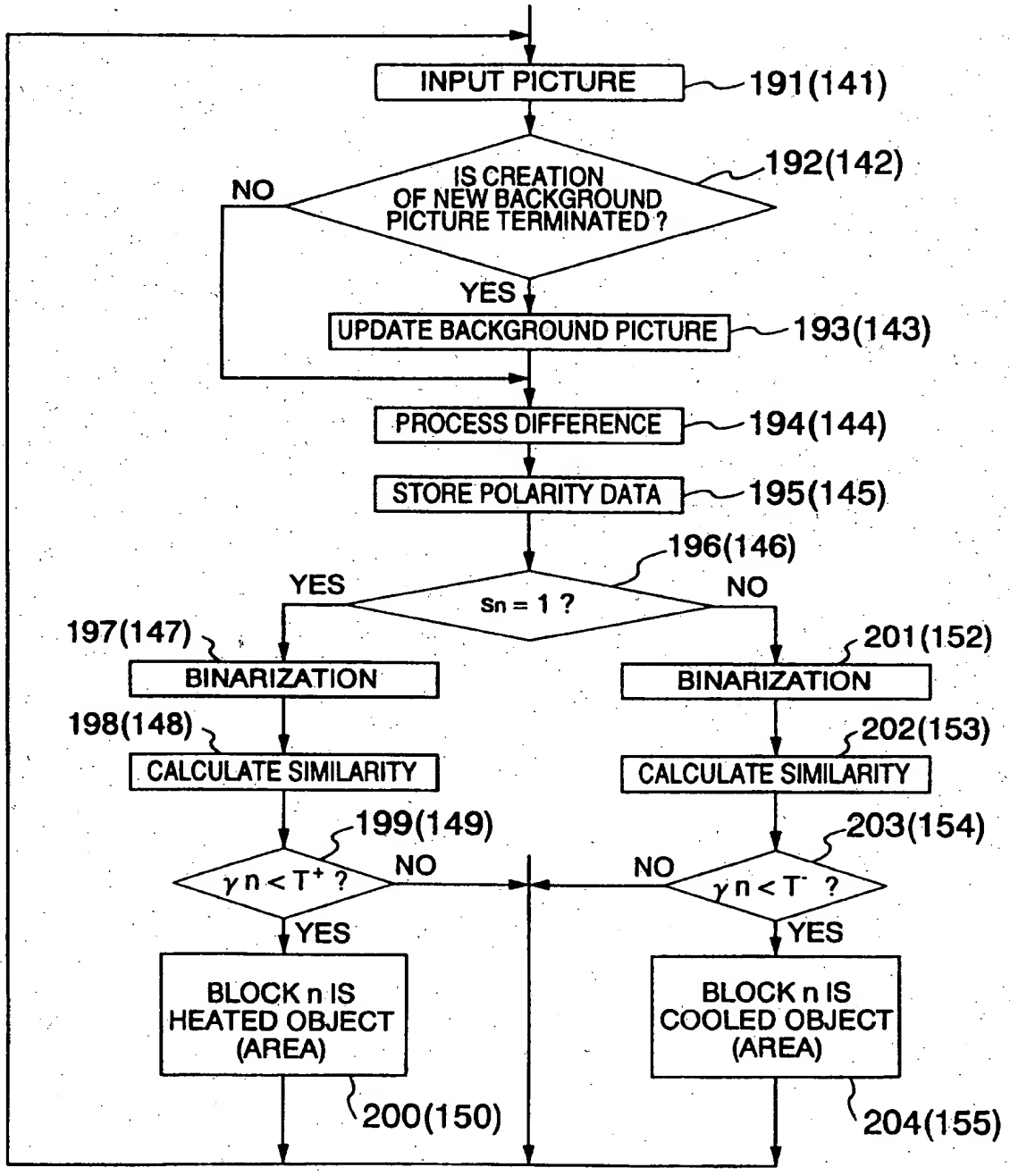

INPUT PICTURE
191(141)
IS CREATION OF NEW BACKGROUND PICTURE TERMINATED ?
192(142)
NO
YES
UPDATE BACKGROUND PICTURE
193(143)
PROCESS DIFFERENCE
194(144)
STORE POLARITY DATA
195(145)
196(146)
Sn = 1 ?
YES
NO
197(147)
BINARIZATION
198(148)
CALCULATE SIMILARITY
199(149)
γn < T+ ?
NO
YES
BLOCK n IS HEATED OBJECT (AREA)
200(150)
201(152)
BINARIZATION
202(153)
CALCULATE SIMILARITY
203(154)
γn < T- ?
NO
YES
BLOCK n IS COOLED OBJECT (AREA)
204(155)

FIG.18

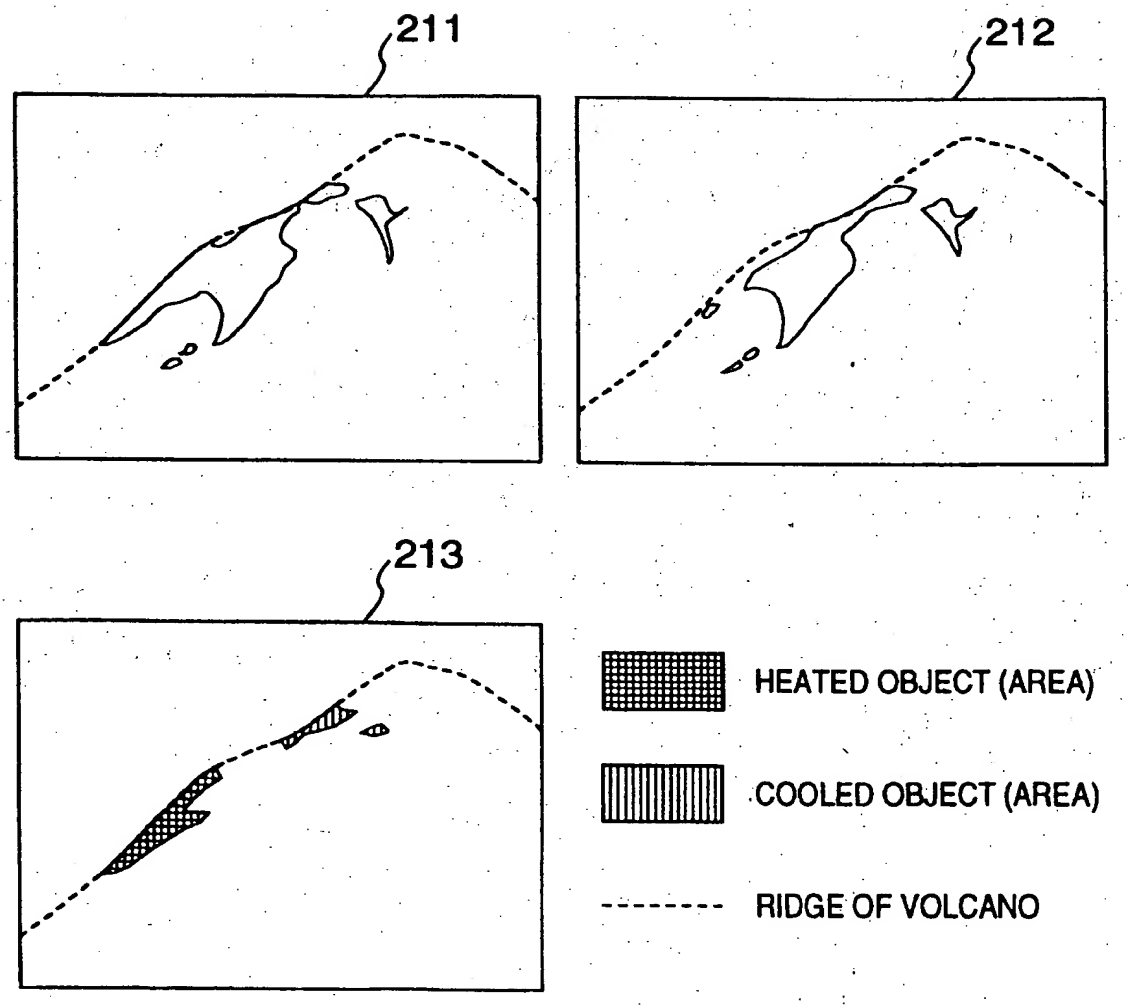

211
212
213
HEATED OBJECT (AREA)
COOLED OBJECT (AREA)
RIDGE OF VOLCANO

FIG.19

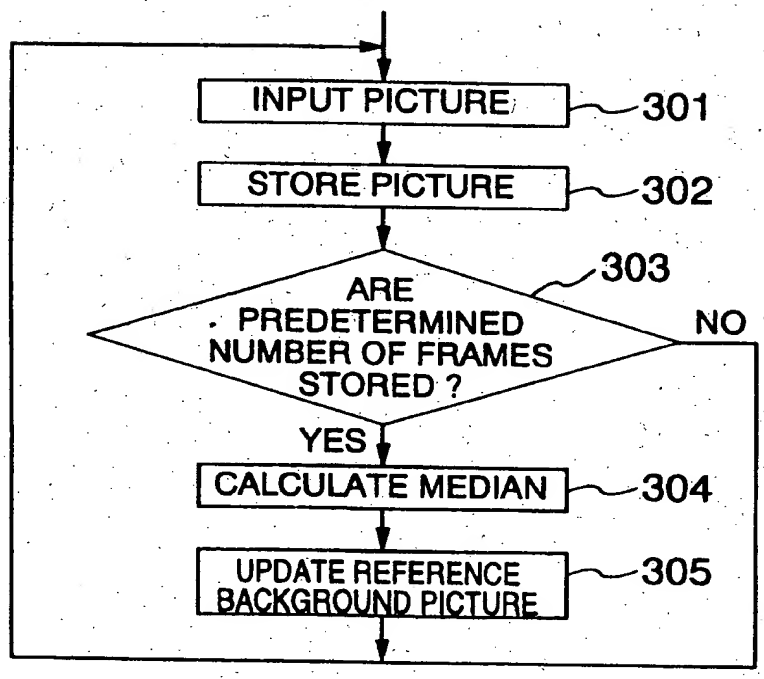

INPUT PICTURE
301
STORE PICTURE
302
ARE PREDETERMINED NUMBER OF FRAMES STORED ?
303
NO
YES
CALCULATE MEDIAN
304
UPDATE REFERENCE BACKGROUND PICTURE
305

FIG.20A

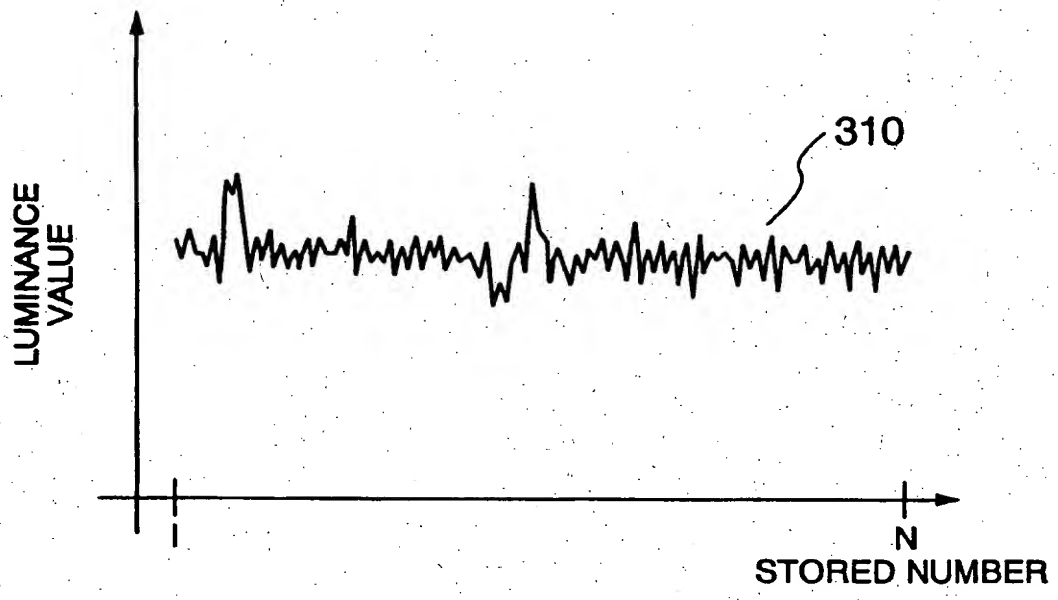


FIG.20B

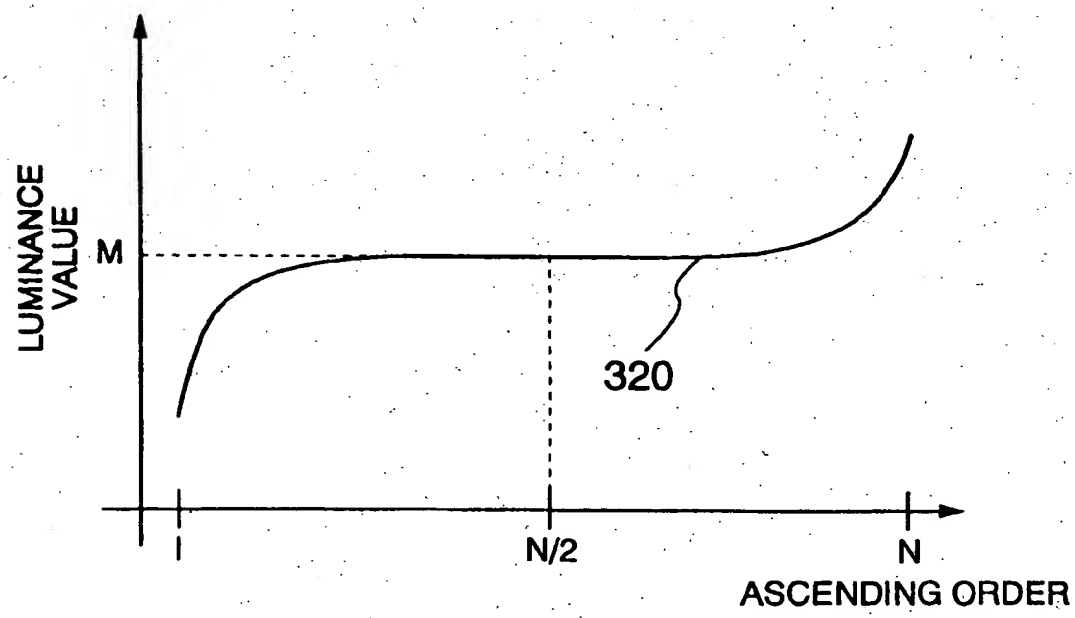


FIG.21

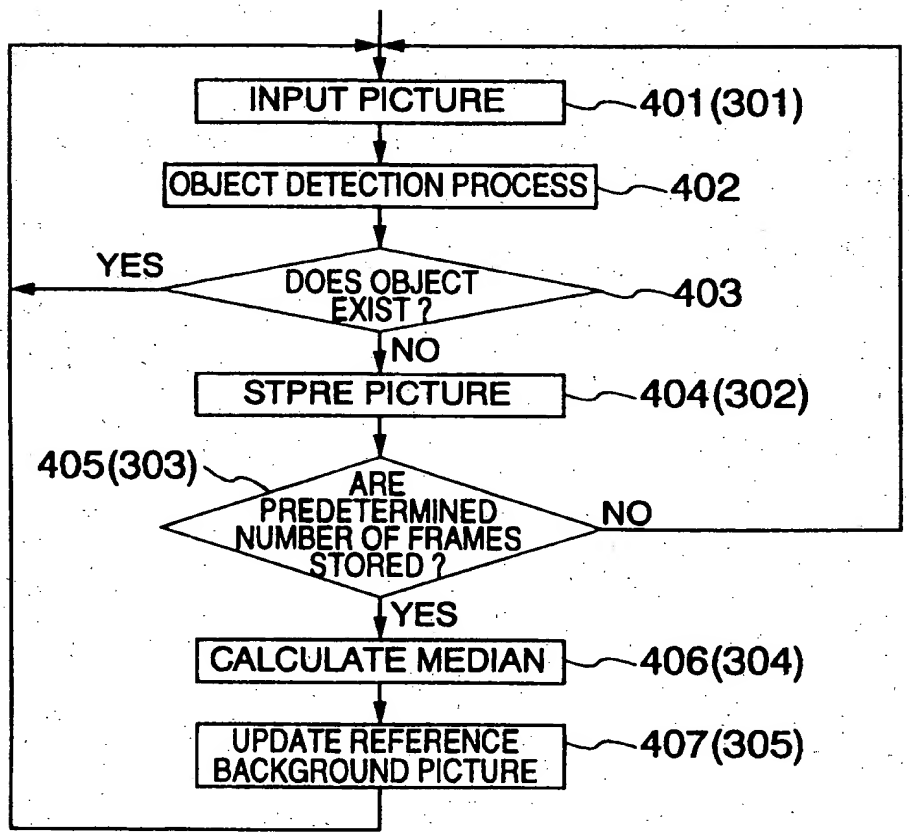

INPUT PICTURE
401(301)
OBJECT DETECTION PROCESS
402
YES
DOES OBJECT EXIST ?
403
NO
STPRE PICTURE
404(302)
405(303)
ARE PREDETERMINED NUMBER OF FRAMES STORED ?
NO
YES
CALCULATE MEDIAN
406(304)
UPDATE REFERENCE BACKGROUND PICTURE
407(305)

FIG.22

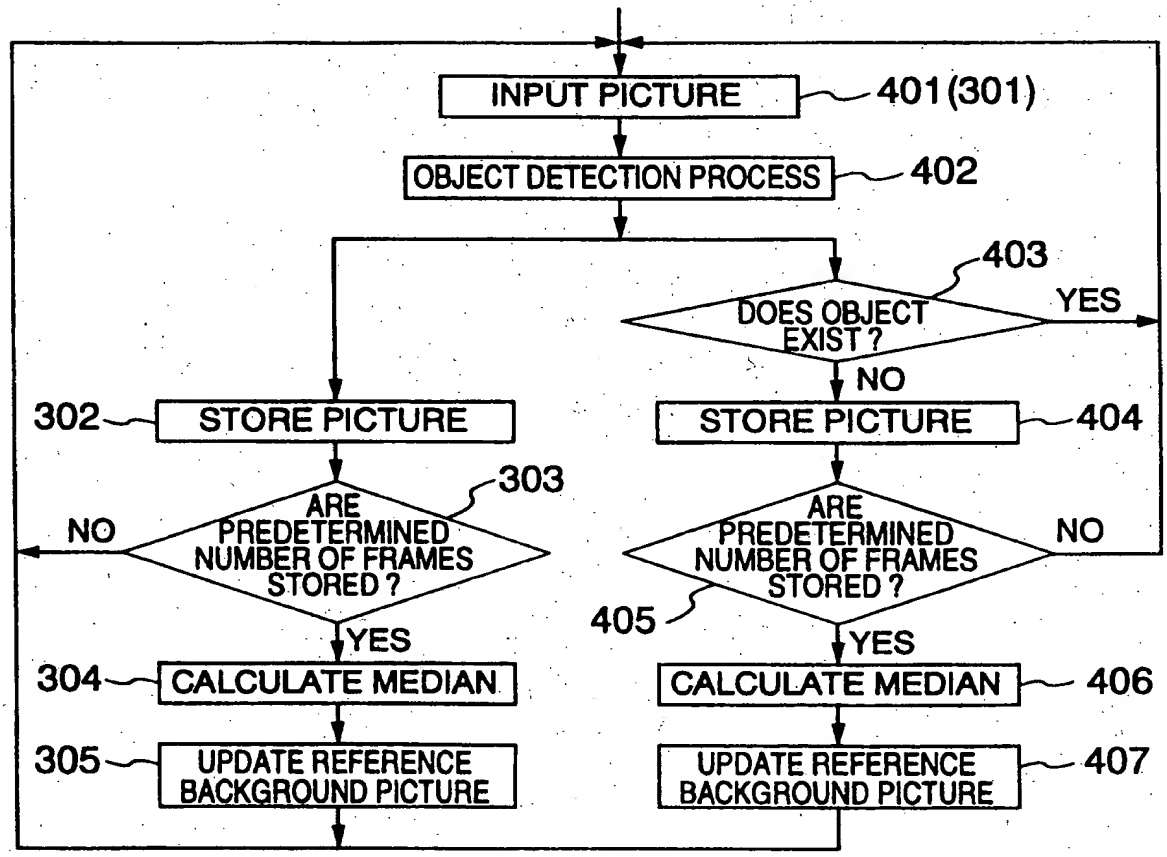

INPUT PICTURE
401(301)
OBJECT DETECTION PROCESS
402
403
DOES OBJECT EXIST ?
YES
NO
302
STORE PICTURE
STORE PICTURE
404
303
ARE PREDETERMINED NUMBER OF FRAMES STORED ?
NO
ARE PREDETERMINED NUMBER OF FRAMES STORED ?
NO
405
YES
YES
304
CALCULATE MEDIAN
CALCULATE MEDIAN
406
305
UPDATE REFERENCE BACKGROUND PICTURE
UPDATE REFERENCE BACKGROUND PICTURE
407

FIG.23

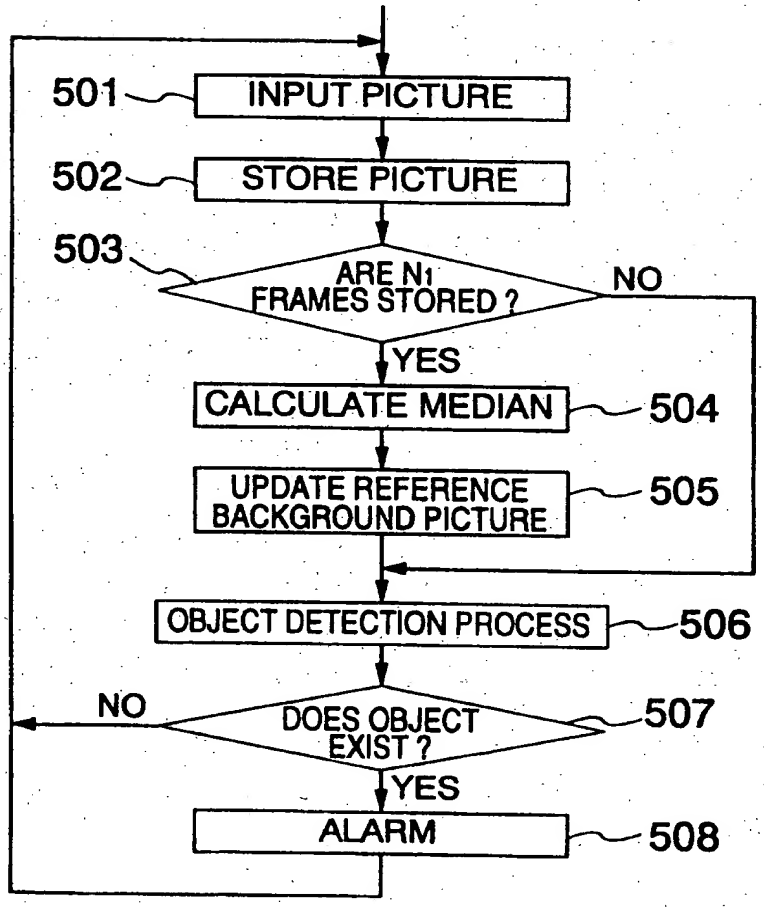
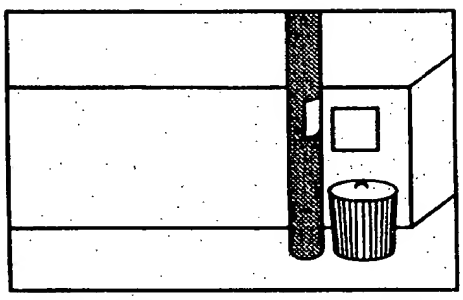
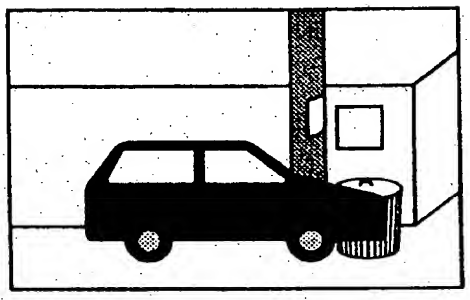

501
INPUT PICTURE
502
STORE PICTURE
503
ARE N1 FRAMES STORED ?
NO
YES
CALCULATE MEDIAN
504
UPDATE REFERENCE BACKGROUND PICTURE
505
OBJECT DETECTION PROCESS
506
NO
DOES OBJECT EXIST ?
507
YES
ALARM
508

FIG.24A



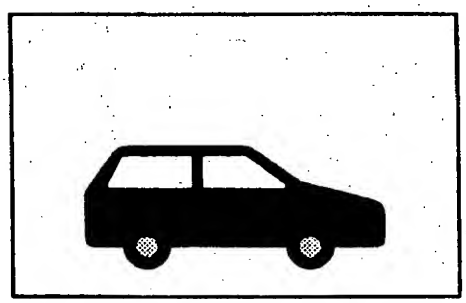
BACKGROUND PICTURE

FIG.24B



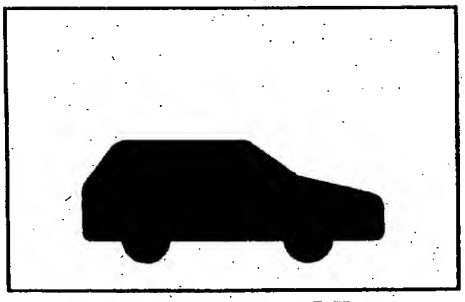
INPUT PICTURE

FIG.24C



DIFFERENCE PROCESSED PICTURE

FIG.24D



BINARY IMAGE

FIG.25

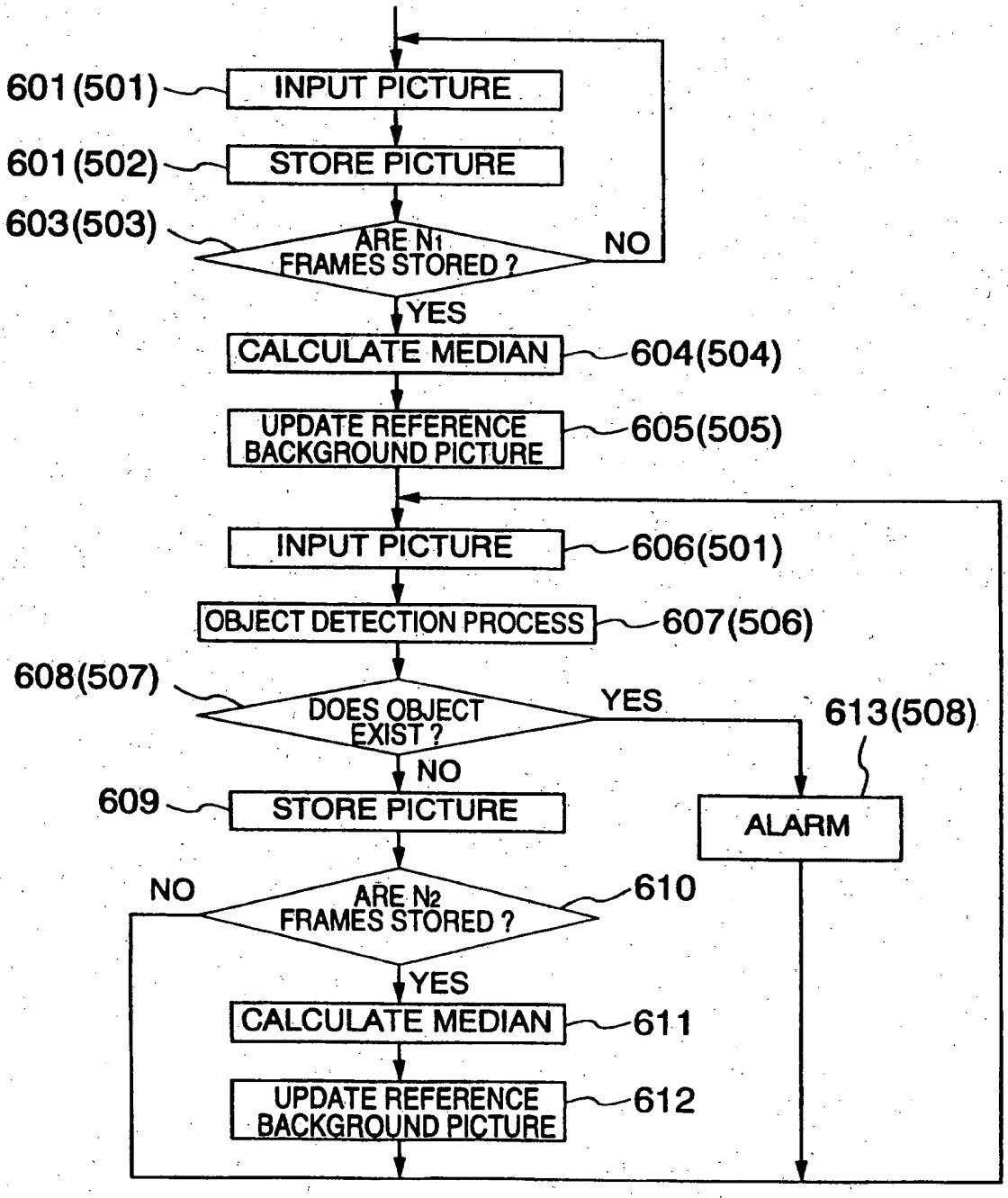
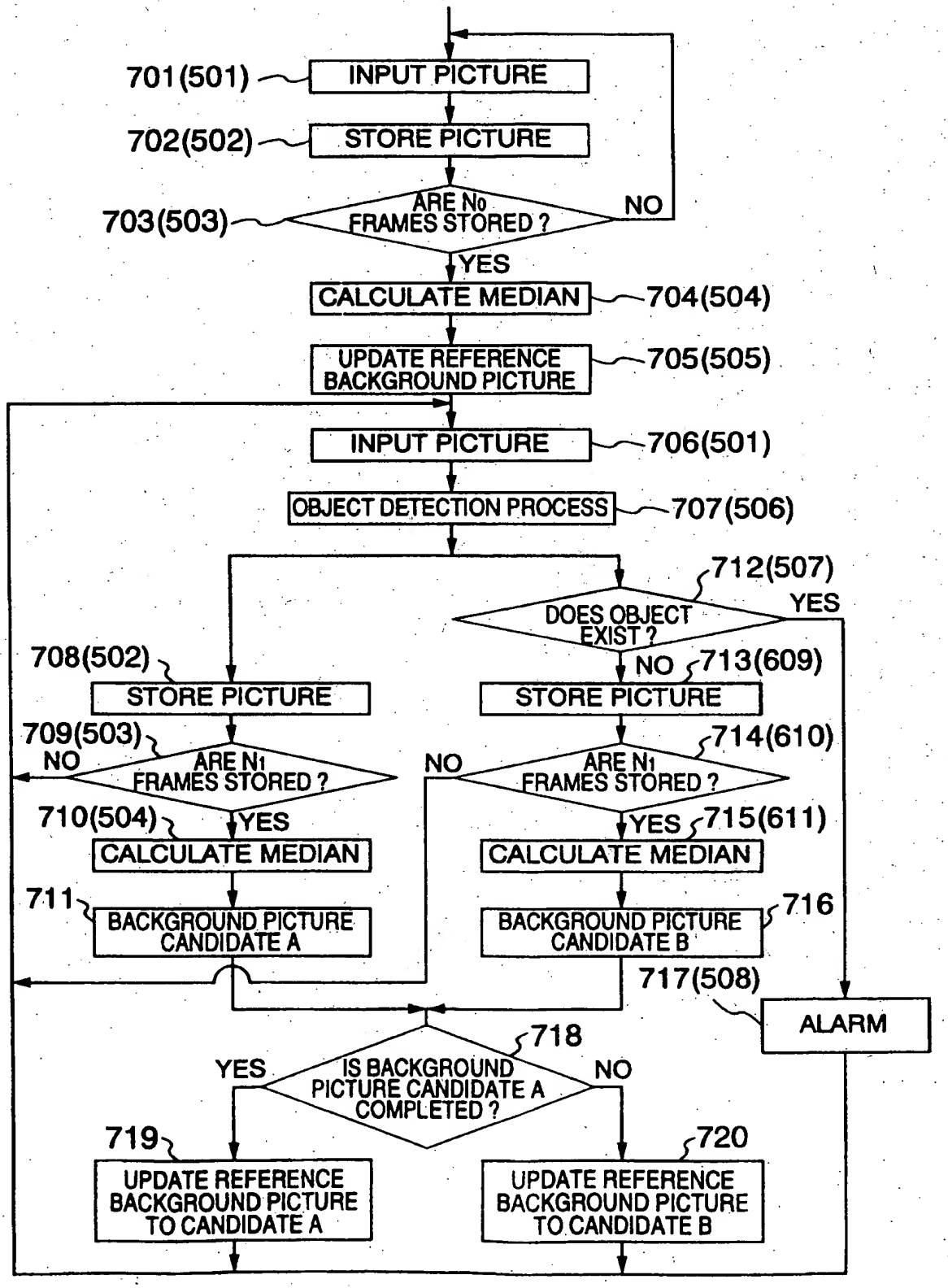

601(501)
INPUT PICTURE
601(502)
STORE PICTURE
603(503)
ARE N1 FRAMES STORED ?
NO
YES
CALCULATE MEDIAN
604(504)
UPDATE REFERENCE BACKGROUND PICTURE
605(505)
INPUT PICTURE
606(501)
OBJECT DETECTION PROCESS
607(506)
608(507)
DOES OBJECT EXIST ?
YES
613(508)
NO
609
STORE PICTURE
ALARM
NO
ARE N2 FRAMES STORED ?
610
YES
CALCULATE MEDIAN
611
UPDATE REFERENCE BACKGROUND PICTURE
612

FIG.26


701(501)
INPUT PICTURE
702(502)
STORE PICTURE
703(503)
ARE N0 FRAMES STORED ?
NO
YES
CALCULATE MEDIAN
704(504)
UPDATE REFERENCE BACKGROUND PICTURE
705(505)
INPUT PICTURE
706(501)
OBJECT DETECTION PROCESS
707(506)
712(507)
DOES OBJECT EXIST ?
YES
NO
708(502)
STORE PICTURE
713(609)
STORE PICTURE
709(503)
ARE N1 FRAMES STORED ?
NO
714(610)
ARE N1 FRAMES STORED ?
NO
YES
710(504)
CALCULATE MEDIAN
YES
715(611)
CALCULATE MEDIAN
711
BACKGROUND PICTURE CANDIDATE A
716
BACKGROUND PICTURE CANDIDATE B
717(508)
ALARM
718
IS BACKGROUND PICTURE CANDIDATE A COMPLETED ?
YES
NO
719
UPDATE REFERENCE BACKGROUND PICTURE TO CANDIDATE A
720
UPDATE REFERENCE BACKGROUND PICTURE TO CANDIDATE B

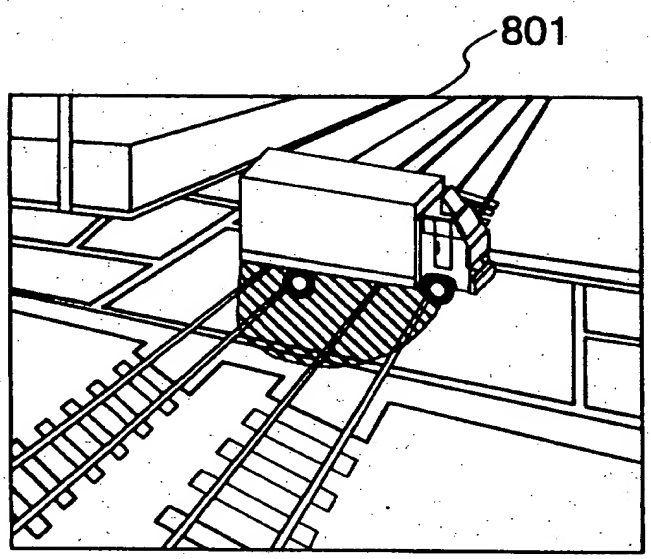
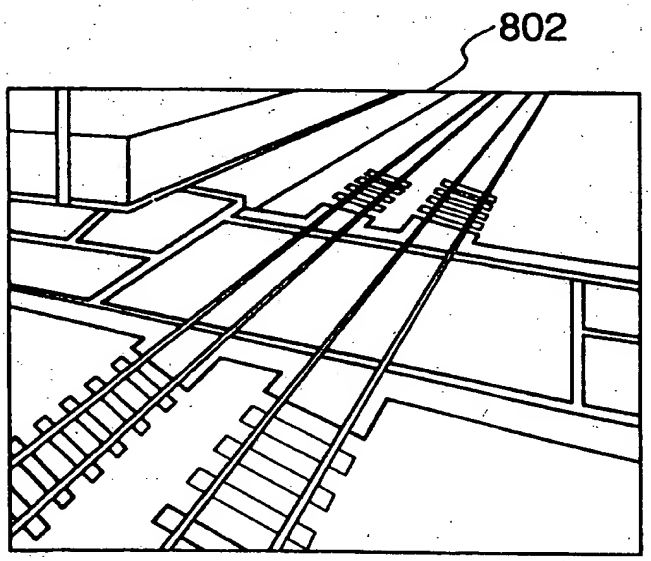



FIG.27